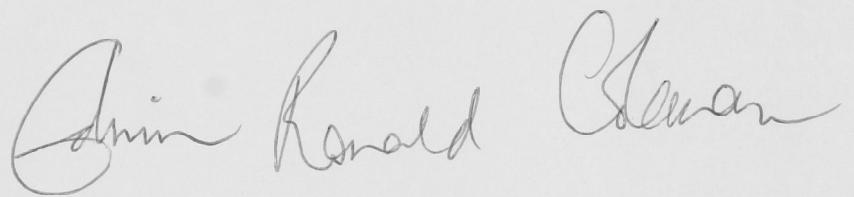


ACKNOWLEDGEMENT

This thesis is my own work and I have acknowledged all the sources on which I have drawn in the bibliography.

I would like to thank Peter Roeper and Quentin Gibson, my successive supervisors, for all their help.

Edwin Ronald Coleman

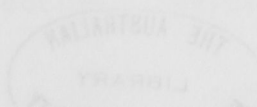
A handwritten signature in cursive script that reads "Edwin Ronald Coleman". The signature is written in dark ink and is positioned below the printed name.

RATIONALITY AND EVOLUTION

by EDWIN RONALD COLEMAN

A thesis submitted for the degree of Master of
Arts in the Australian National University

August, 1977.



Contents

Part One : Rationality

Chapter One : Rationality And Frameworks

- 1.1 Reason and science.
- 1.2 The argument for frameworks and the logical core of part one.
- 1.3 The Kuhn/Popper dispute.
- 1.4 The argument from frameworks.
- 1.5 The notion of framework.
- 1.6 The argument against frameworks.
 - (A) frameworks as total.
 - (B) frameworks as partial.
- 1.7 Rules.

Chapter Two : Rationality And Conceptual Revision

- 2.0 Preamble.
- 2.1 Rules of discourse.
- 2.2 Bennett's thesis.
- 2.3 Applying concepts.
- 2.4 Conceptual change.

Chapter Three : Rationality And Logical Revision

- 3.0 Connecting summary.
- 3.1 Examples of revisions of logic.
- 3.2 First objection : revision is pretended.
- 3.3 Second objection : revision is trivial.
- 3.4 Third objection : revision is irrational.
- 3.5 Fourth objection : revision can only be peripheral.

Part Two : Evolution

Chapter 4 : Naive Evolutionary Epistemology

- 4.1 Some motivations for evolutionary epistemology.
- 4.2 Naive evolutionary epistemology in the literature.
- 4.3 Pressing the analogy.

Chapter 5 : The Logical Form Of Evolutionary Explanation

- 5.1 Introduction and remarks on extant literature.
- 5.2 The Evolutionary Explanation.
- 5.3 Selection explanation.
- 5.4 Functional explanation and evolutionary explanation.

Chapter 6 : Evolutionary Rationality

- 6.1 A rational explanation is an evolutionary explanation.
- 6.2 The duality of rational life.
- 6.3 Ontology of rational life.
- 6.4 Causal phases of rational life : acquisitions.
- 6.5 Cybernetic phases of rational life : repudiations.
- 6.6 Statistics of rational selection.
- 6.7 The cybernetic content of theories.
- 6.8 The evolution of rational life : two examples.

PART ONE

RATIONALITY

CHAPTER ONE

RATIONALITY AND FRAMEWORKS

1.1 Reason And Science

Man has long been supposed distinguished from the animals by his reason¹; on what reason consists in, there has been something less than consensus. There was, though, until quite recently, consensus that reason could be seen exercised in science and mathematics in a particularly pure and perspicuous way. In recent philosophy of science, however, this view has been under strong challenge from Kuhn² and Feyerabend³ and their disciples. Indeed, doubt has been cast on the idea that science is fundamentally rational by the use of detailed historical research to show that certain key scientific changes are not rational. This is supported by general arguments as to why such episodes should be expected. The doctrine may be called the thesis of incommensurable frameworks. It is a somewhat more precise version of the doctrine which has been around for some time, underpinning relativism, traceable to Wittgenstein's later work⁴.

In this essay I propose that work such as that of Kuhn and Feyerabend shows not that science is irrational au fond but that the common notion of reason is wrong. More specifically, I shall argue for a third alternative between an absolutist conception of reason as an unchanging and unchangeable Platonic given - which is refuted by Kuhn and Feyerabend - and a stultifying relativism about reason which is commonly taken as the only alternative. My thesis will be

1. cf. Aristotle, (DA), 428^a 23

2. (SSR), (PSSR) and (RC)

3. (AM), and a series of earlier articles this book sums up.

4. cf. Winch (ISS), Wittgenstein (OC, PI)

that evolutionary naturalism provides an account of human reason as a faculty subject to change, but to change which is not arbitrary and directionless. In short, I propose that we conceive rationality, like other human characteristics, as an evolutionary product and process.

This positive suggestion is the subject of part 2 of the essay. In part 1 I shall criticize what I take to be the prevailing concept of rationality, locating its weaknesses in various different forms of the assumption of reason's unchangingness; in each case I find that a certain characteristic logical error is involved, which I shall dub the Urmutter fallacy.

In the first chapter I set out what I take the essential Kuhn/Feyerabend argument from frameworks to be. I then criticize the systematic ambiguity in the concept of framework on which this argument depends. In chapters 2 and 3 of part 1, I further criticize the notion of reason which the argument from frameworks employs, in its implications of the fixity of concepts and the fixity of logic respectively. Throughout this whole discussion the role of rules in rationality is under consideration.

A brief characterisation of my attitude in part 1 is this : the rules which govern discourse and thought can be changed rationally as well as followed.

1.2 The argument for frameworks and the logical core of Part 1.

The following argument is quite basic to the whole of part 1 of this essay.

- (A) For an entity to be rational, it must be able to give reasons and to assess reasons.
- (B) Consequently, a prerequisite of reason is a linguistic community.
- (C) Giving reasons involves the exercise of concepts.
- (D) Exercising concepts requires following rules.
- (E) Assessing reasons requires the use of logic.
- (F) Using logic requires following rules.
- (G) Hence, reason presupposes a minimal conceptual endowment and a minimal logical competence.
- (H) Reason requires following rules.

I accept, for present purposes, all of this argument, although many reservations are in order.

A and B depend upon something like the private language argument, of which I am not totally convinced; but part 2 will make such doubts irrelevant to my main points.

C and D are only as acceptable as the content given to the notions of 'concept' and 'rule' - which will be discussed in detail below.

E and F are only as acceptable as the notion of logic involved, and this will be discussed in chapter 3.

Nonetheless, I accept the conclusion H. But a fundamental point, which constitutes the logical core of part one of this essay is that this argument does not commit us to accepting that reason presupposes either some specific set of ineluctable concepts, or some specific unavoidable logical apparatus, or some particular rules.

To treat these points together, although it seems unexceptionable to assert that

(1) "X exercises rationality at time T " entails

"there are certain rules (concepts, logic) to which X adheres at time T "

it does not seem to be true, and in any case it certainly does not follow from (1), that

(2) "there are certain rules (R) (concepts C, logic L) such that if X exercises rationality at time T, X adheres to R (C,L) at time T".

To suppose that (2) follows from (1) is to commit a logical fallacy, which I call the Urmutter fallacy, for its form is that of this inference:

"Everyone has some mother, therefore, someone is the mother of all".

It is as though one were to argue from the undoubted truth of

(3) "X exercises his arms at T" entails

"there are cells which X contains at T"

(the cells compose X's arms)

to the undoubted falsity of

(4) "There are certain cells C such that if X exercises his arms at T, X contains C at T"

for we are not all one organism!

The argument

$$(R(x,t) \rightarrow (\forall c) (A(x,t,c)))$$

$$\rightarrow (\forall c) (R(x,t) \rightarrow A(x,t,c))$$

is invalid

Half of my aim in part 1 is to suggest that a conception of reason which is prevalent and underlies a horrified reaction to Kuhn & Feyerabend, commits this fallacy. The other half is to oppose the opposite error, which is to espouse this denial, while simultaneously denying that it can be true that

- (5) "At T, X has concepts C_1 , and logic L_1 , at T^* , later, X has concepts C_2 and logic L_2 . $C_1 \neq C_2$, $L_1 \neq L_2$ yet X proceeded rationally from T to T^* ,"

The intermediate thesis which can assure us of the truth (sometimes) of (5) is

- (6) There are concepts and logic C_3, L_3 which X had from T to T^* inclusive.

But (6) does not commit us to (2).

1.3 The Kuhn/Popper dispute

In "Normal Science and Its Dangers"⁵, Popper declares that the fundamental difference between him and Kuhn concerns the "logical thesis" (sic) that

"rationality depends upon something like a common language and a common set of assumptions".⁶

5. in (CGK) p.56

6. ibid p.56. Ironically, Kuhn accuses Popper of the same : (CGK) p.234!

Popper declares this to be a mistake, for

"we are prisoners caught in the framework of our theories
(only) in a Pickwickian sense; if we try we can break out
of our framework at any time ... the central point is that a
critical discussion and a comparison of the various frameworks
is always possible ..."⁷

Kuhn does not accept this :

"if frameworks are the prerequisite of research, their grip on
the mind is not merely 'Pickwickian' ... to be simultaneously
essential and freely dispensable is very nearly a contradiction
in terms. My critics become incoherent when they embrace it."⁸

This passage enshrines a version of the characteristic error called
the Urmutter fallacy. Kuhn is implicitly using the phrases
"simultaneously essential" and "freely dispensable" to refer to some
definite unique framework. The Popperian point is that although some
framework must be employed at any given point in time by a specific
thinker, no specific framework is essential - each is freely dispensable.

Feyerabend also rejects Popper's views : in defending his thesis that

"there are frameworks of thought which are incommensurable".⁹
he rejects the claim of Popper I quoted above by asserting that
translations into a common language to conduct the discussion always
"do violence" to one language concerned.¹⁰

7. ibid p. 56.

8. ibid p. 242.

9. (AM) p.270.

10. ibid p.272, n.130

The supposed force of the claim that scientific thought must depend on a framework may be that it explains the normative character of the thought, the unanimity among the competent, in terms of logical compulsion.

Thus the inference,

- (i) theory A is simpler than theory B
 - (ii) the simpler of two theories is preferable to the other
- so (iii) A is preferable to B

is valid, so that anyone accepting the premises must accept the conclusion. Thus, a shared framework includes shared rules, which control inference - logical rules. Kuhn seems to have this in mind, for he speaks thus of cases when the compulsive unanimity is lacking :-

"only if the two (disputants) discover ... that they differ about the meaning or applicability of a stipulated rule, that their prior agreement does not suffice for proof, does the ensuing debate resemble what inevitably occurs in science."¹¹

however,

"What I am denying is neither the existence of good reasons nor that these reasons are of the sort usually described. I am, however, insisting that such reasons constitute values to be used in making choices rather than rules of choice."¹²

11. (CGK) p.261

12. *ibid.* Note that Kuhn is not here raising the question of the revisability of logic : it is not the inference but the premises which is in dispute.

Kuhn is here attempting to deny that he regards "deep" theory-choice as irrational, by calling the application of values like simplicity, "giving reasons". He clearly implies that most compelling reasons are rules, specifically rules of choice.

At this point Kuhn seems to be identifying all rules with logical rules, yet elsewhere¹³ he allows for at least four kinds. The point seems to be that the kind of debate he is concerned with is not about the propriety of an inference but "about premises".¹⁴ Why rules should not be required in such debates is not made clear. Two points must be cleared up here:

Firstly, what kind or kinds of rules does a framework involve, and what are their roles in reasoning? Secondly, how far does rationality depend on different kinds of rules and their interplay? These matters will be discussed below in chapter 2.

The absence of shared rules, to which Kuhn is pointing in the last quotation is also one of Feyerabend's major claims:

"There is not a single rule that remains valid under all circumstances ..."¹⁵

"None of the methods which Carnap, Hempel, Nagel, Popper or even Lakatos want to use for rationalizing scientific changes can be applied ... what remains are our subjective wishes"¹⁶

This claim is directly opposed to the professed aim of Lakatos, which is just to find

13. (PSSR), P.39ff.

14. (SSR) P.40 et. seq.

15. (AM) p.213.

16. *ibid.* p.285

"universal demarcation criteria ... (so that) all major changes in science can be explained using the same single criterion of scientific merit ... the methodology of scientific research programs is a new demarcationalist methodology (i.e. a universal definition of progress) ..."¹⁷

My position is somewhere between these two : I agree with Feyerabend that there is no universal rule, but I also agree with Lakatos' aim insofar as it can be achieved : namely, in rationalizing scientific change. There is a fundamental difference in seeking a rule to cover all known cases (this can always be done since they are finitely many, though an interesting one may be hard to find), and in seeking a rule to cover all possible cases. The latter I do not think possible : if it could be found, it would legislate for future science in an absolute way. But, I shall argue below, every rule is revisable (ch. 3).

Eventually I shall be agreeing much more with Kuhn than ~~with Feyerabend~~ (the former having mutilated, at least, his original position by a thousand qualifications), but at the present point I will treat them as propounding the same argument, which I shall shortly set out. The trouble with Kuhn's position is that even though he calls the process of paradigm-change rational, his characterization is so weak that he can give no content to the idea that progress really occurs through revolutions (beyond the rewriting of history by the victors). This simply ignores the fact that although, as is becoming very well-known, Newton's theory contradicts Galileo's, it remains the case (as, in effect, Popper pointed out in the same article¹⁸ in which the contradiction was first brought out), that the relation

17. Lakatos, (LZ) pp. 364, 368.

18. Popper, (AS)

between them is asymmetric - Newton's explains why Galileo's is wrong, but Galileo's does not explain why Newton's is. This is progress, this does make it rational in a stronger sense than Kuhn admits, for preferring the Newtonian theory to the Galilean.

1.4 The Argument From Frameworks

I take the argument jointly proposed by Kuhn and Feyerabend to be this :-

- (P1) There are genuinely different frameworks. (premise)
- (P2) A rational decision requires the use of rules. (premise)
- (P3) The employment of rules must be part of the employment of a framework. (premise)

From (P2) and (P3) is deduced

- (C4) Any rational decision is made employing a framework.

From (C4) and the premise (P1) it follows that

- (C5) Giving up one framework for a genuinely different one cannot be a rational decision.

This general relativistic thesis is applied to the case of science thus

- (P6) For some ("deep") scientific theories, acceptance commits one to some specific framework. (premise)
- (P7) The frameworks involved by some pairs of such theories which are in the relation of replaced and replacing theories, are genuinely different. (premise)
- (I call such a pair a succession pair)

From (P6) and (P7) we may deduce

(C8) For some succession pairs of theories, choice between them involves giving up one framework for a genuinely different one.

(C8) now combines with the thesis (C5) to yield

(C9) The choice between some succession pairs of theories is not rational.

I accept premisses (P2) and (P3), and a qualified (P1) - qualified as to the meaning of "genuinely different", which I examine below. I do not accept (C5) however, and I shall argue against it thus : there are two senses in which 'framework' may be used in this argument, which I will spell out in 1.4, a 'partial' and a 'total' sense.

If the 'partial' sense is meant, (C5) does not follow from (P1) and (C4); if the 'total' sense is meant, then (P1) is false and (C5) is nugatory - what it calls irrational can never happen!

My further argument against (C9) consists in pointing out that for (P6) to be true, the 'partial' sense of 'framework' must be used; but in this sense of 'framework' (P7) is false.

Moreover, the second premise will be closely examined in chapter two. The truth of (2) will then be seen to be dependent on a construal which requires a 'partial' notion of framework. The premise (3) is discussed in the next section of this chapter, as are premisses (6) and (7). In chapter 3 I shall discuss a thesis which must be true if (1) is to be true on the 'total' interpretation of framework - the fixity of logic - and find it wanting. That is, I shall argue that logic is reviseable, and so there are different frameworks, all right, but not "genuinely" different - for there cannot be "genuinely" different logics for them.

1.5 The Notion Of Framework

Just what is a framework, however? There are many more or less similar notions abroad which might be meant. Form-of-life,¹⁹ system of absolute presuppositions,²⁰ paradigm²¹ (in one sense!), the tacit dimension²², conceptual scheme²³ and so on and on. They all derive from Kant's apriori elements of experience by some kind of relativisation.

The clearest exposition of such a notion is Körner's²⁴, under the name categorial framework, and it is also the most explicitly Kantian. The underlying thesis, for which Körner does not argue, is the thesis (G) of 1.2, that reason entails a minimal conceptual endowment and a minimal logical competence.

Körner's explication of (G) is this : all thinking requires a classification into maximal kinds, and this analyses out to a set of concepts articulated by certain principles, constitutive and individuating. Since these take the form of logical relations between concepts, mainly entailment and disjunction, clearly some logical principles, at least sufficient to characterise what is to count as entailment or disjunction, are also required.

Körner is rather vague about what counts as a logic, giving the impression that a logic may be identified by (perhaps even with) a calculus which can represent it. This somewhat narrow conception of logic, as the minimal requirement for characterizing the logical concepts used, is discussed below in Chapter 3.

19. Wittgenstein (PI), (OC)
 20. Collingwood (EM)
 21. Kuhn (SSR)
 22. Polanyi (PK)

23. Davidson (VICS)
 24. Körner, (CF)
 25. Bennett, (R)

Now the main problem about frameworks for present purposes is, are they total? That is, are they all-encompassing - is everything grist to their mill? This is not a way of begging the question, a way of eliminating the supposed impossibility of talking about phlogiston in atomic terms; rather I am raising the question whether atomic terms purport to treat of dreams or propositions, sins or human rights.

Körner's concept of a categorization is clearly meant to be total in this way, for he says "the notions of an object and of a characteristic cannot be defined independently"²⁶ and "it will be convenient to reserve the term 'categorization' for the higher levels of a total classification."²⁷

On this notion, every object which the user of a framework "discerns", (Körner's term,) must fall into one of the maximal kinds of the categorization.

Let us take it for present purposes that a total classification is one such that anything its user discerns is in some class (perhaps more than one); let the test of discernment be the possibility of linguistic reference. Then a categorization in Körner's sense must be referentially complete. Let us call the constraints on a person's discourse beyond this the logical competence of that person. Then I shall say that a framework is dialectically complete if it entails the logical competence of the person in question. So we shall say that a framework which someone is using is complete if it is both referentially and dialectically complete. It is then the

26. Körner, (CF), p.1

27. *ibid.* p.4. The apparent weakening to "all objects or all objects of a natural class" in the next sentence refers to the 'secondary phases' of the categorization.

indispensable part of a person's cognitive competence. Thus neither the concept "blue skies", nor Pythagoras' theorem, would be part of my present framework, since they can be recovered from the rest of my present cognitive competence. The framework is the core remaining when such dependent elements are stripped away. The assumption that there is a unique identifiable core will not be criticized at this point, but cf. ch.3.

By a framework in the partial sense, I shall mean a subset of a total framework; typically, it will be the framework obtained from a person's cognitive competence concerning some domain of experience.

I shall argue in the next section that a theory is always (at most) a partial, and not a total, framework.

1.6 The Argument Against Frameworks

Clearly, any alleged framework is either total or partial in the sense I have indicated.

I will now give some arguments to show that on either interpretation, at least two of the premises in the argument 1.4 are false. I shall draw the obvious conclusion that the charge of irrationality is, as the Scots have it, 'not proven'.

(A) Frameworks as total.

(A1) If frameworks are total, then (P1) is false - at least, we could never be in a position of having evidence for the truth of it. Applying Ockham's razor then requires its rejection. Evidence for (P1) would need to be either a description of two

allegedly different frameworks, in some strong sense of different, or else phenomena which could have no better explanation than one assuming the truth of (P1).

A description of two frameworks which would be comprehensible would have to interpret one into the other, or else both into a third. In the former case they are not different in any epistemologically dangerous sense, in the latter case they can only be total if all three are mutually interpretable, which is the former case all over again. This dilemma arises from the fact that the interpretation of F_1 in F_3 must be either F_3 or a "subframework" of it. But the latter case is ruled out by the 'totality' reading of frameworks, since the classes of objects involved by F_3 but not by the interpretation of F_1 must be discernible by an acceptor of F_1 , since F_3 shows how a coherent extension of F_1 could be constructed. The same argument applies to the other of the allegedly 'different' frameworks.

What about the other possible kind of evidence, namely puzzling phenomena whose explanation requires (P1)? I think that this rather vaguely specified possibility must in fact succumb to a similar argument. The phenomena in question would have to consist in highly divergent behaviour by typical members of the different groups of people in what are ostensibly identical circumstances. Now this is an empirically well-evidenced phenomenon, and one can certainly deduce that some of the beliefs held by the respective typical members of the two groups must differ. But what must the phenomena be like for (P1) to be concluded? It would surely be required that in no circumstances ostensibly the same, was the behaviour the same. I am quite sure that there are no

ethnographies to that effect; and if there are, we would have to ask whether there was any reason to think that the more mysterious group consisted of rational beings having beliefs at all. They would presumably be apparently human beings who did not eat when hungry, talk about the things around them, or have a word for 'human being' or 'person', or for any of the other obvious furniture of the world common to all men. It is obvious that this is just a fantasy, that languages used by humans (or even near-humans) must have a great deal of semantic overlap just because of the inevitable identity of the more mundane aspects of existence (food, sex, death, etc.), consequent upon the physiological unity of mankind.

- (A2) Still regarding frameworks as needing to be total, it seems that (P_6) must also be false. That is, no theory is, or could be a framework, (even in the weak sense that acceptance of the theory commits one to a specific framework).

Firstly, none of the standard examples can count as a framework on the total view. Even Feyerabend's insistence that there is no theory-neutral observation language counts against (P_6) since he wants us to be alive to the way the 'data' may be tainted by older ideologies (AM, ch.5) and insofar not by the theories under examination. In the case of the clash between Aristotelian and Galilean physics, when for example Kuhn²⁸ is alleging that scientists could see only a "pendulum" or "constrained fall", his own description is 'the scientist who looks at a swinging stone' which has to be sufficiently neutral to the theories which

28. Kuhn, (SSR) p.128.

impregnate the previous two phrases, to allow the reader some independent access to what he is talking about. The point is quite general and might be connected with the argument under A1 through the self-defeating character of the attempt to describe the alleged cases.

In fact, I think, no scientific theory can be a framework. Even if we were to grant the reductionist thesis, that all theories (psychological, economic, etc. ...) will one day be one, even if we grant what is patently untrue that that theory will be, say, existing physical theory, even then existing physical theory would not count as a framework because users of it do not acknowledge into which maximal kinds of physical theory fall, say, dreams (physical object or property?) or money - precisely because the theoretical reductions which would allow this do not yet exist. A doctrinaire materialist might have a dogmatic attitude toward this theory, declaring that atoms are atoms and everything else is a pattern of atoms - but this is not part of the theory, as is obvious from the existence of non-materialists. In any case, since that theory doesn't mention dreams or money or Pythagoras' theorem, a framework which declares them to be patterns of atoms must be distinguishable from it; and the belief that reductions are possible is patently not part of the theory till they are available. A theory plus a certain ideology about it may be a framework, but admitting this distinction will clearly undermine (P7), for it must allow for adherence to the theory without commitment to the ideology (and, a fortiori, the framework). I will return to this point, since the argument 1.4 can only be continued, if this line is taken, on a partial notion of framework.

I am not denying that one may use the term 'framework' in such a way that one might regard atomic theory as a framework. I simply wish to distinguish that use from the use in which the categorization is total. The difference is of extreme importance. It is the difference between a language and a theory which can be stated in that language.

It is obvious that many theories are not languages; but may not languages be theories? A language like English would have to be a congeries of theories, being a product of so many different cultural influences. But still, could not the language of some highly isolated and socially compact tribe be a theory? I do not want to claim this to be inconceivable, but we can ignore the possibility, for two reasons. On the one hand, none of the cases we shall be interested in could be so described. And on the other hand, I require of theories something weaker than even falsifiability (though I demand that of scientific theories), namely rejectability.

A language, however, no matter how misleading in its mistakes as far as terms with no referent (God, etc.) and inaccurate classifications (e.g. stars are planets) are concerned, cannot be rejected - though it can be improved. The theorist must be "outside" the theory in a way which only the superstition of the soul could claim for the speaker's relation to his language.

This point is directly connected with another, namely, that it is greatly exaggerating the situation to say that (some) new theories must be learned like a first language without mediation of the old (as Feyerabend claims on p.292). For Feyerabend's²⁹ own description

29. (AM), p.81 ff.

of Galileo's "propagandist machinations" for persuading his readers to revise their 'natural interpretations', show that they rest on first obtaining their agreement in certain other cases - in the example of the tower argument, agreement that 'motion' is not always obvious.

Yet, could one not learn a theory 'from scratch'? 'The intention to start from scratch, after a complete removal of all natural interpretations, is self-defeating': Feyerabend's reason for this remark is the psychological impossibility - the process of removal must stop before it can be complete. And this seems not only sound, but in direct conflict with the claims of p.272 et seq. that some theories must be and are learned 'from scratch'. In fact there is a logical reason for viewing this claim with some suspicion, which reinforces the suggestion I make above that a theory and a language must be distinct. This is Tarski's demonstration³¹ that no consistent language can contain its own semantics : so, in so far as a theory may be viewed as a (formal) consistent language, just this requires a further component of language in which its application is considered.

This is not to say that a theory is merely a logical calculus or formal system. But it must I think employ some such; and the further essential statements of a theory must be mutually articulated with one another and that formal core, though exactly what this means is not yet quite clear.

(The Hempelian attempt at a statement seems to be petering out.)

30. *ibid* p.76.

31. Tarski, (LSM)

Nevertheless, one might still hope to make out a case for a language which was a theory if no other theories or domains of application were in question (so that the questions of applicability would be settled faute de mieux). Indeed, just such a scenario is the subject of Bennett's book Rationality which I discuss in the next chapter. We shall see that the idea is temporally (in particular, evolutionarily) unviable. There just are other domains and other theories at hand; this is not simply a convenient empirical fact, however, but a logical consequence of the way the world is. Whether the world must be as it is I do not venture to say.

(B) Frameworks as partial

If the argument of 1.4 is taken with the "partial" sense of "framework", then that argument fails in two ways : (C5) does not follow from (P1) and (C4); moreover the truth of (P7) can only be maintained if a very weak sense of "genuinely different" is meant - and the same is true of (P1).

The relativist conclusion (C5) no longer follows from (P1) and (C4), on the partial notion of framework, because the fact that some rules and concepts must be carried through the change (which (P2) requires) no longer conflicts with the differentness of the frameworks being exchanged; since frameworks are not total, what we mean by their being different can only be that they are not identical; it does not preclude their partial overlap (in rules and concepts which survive the change). This possibility has not gone entirely unnoticed of course, but its fatal bearing on the argument from frameworks is obscured by the Urmutter fallacy - the notion that because the common part need not be (and in general will not be) the same from case to case, consequently no case can

be a genuinely rational decision. But I think that the onus is on those who claim that there must be certain specific ubiquitous ingredients of rational decisions to produce some of these alleged epistemological desiderata. Some common suggestions are discussed in the next two chapters - in chapter two, ineluctable concepts, in chapter three, supposed logical necessities such as the law of the excluded middle.

If we take frameworks as partial, we must allow the possibility of two frameworks which have no overlap at all. They might be fairly useless, since they would be logically as well as semantically disjoint, and in particular, none of the frameworks we are interested in is sufficiently logically *modest* to be one of such a pair (cf. chapter three for this), yet we may need to consider the possibility of such disjoint frameworks. Does it not remove my objection to the inference from (P1) and (C4) to (C5)? It does indeed, but at some cost. For either the framework first employed is effectively total for the holder, or it is not. If it is, then there is no sense to the description that a change is made - it is a metamorphosis as radical and inexplicable as that in Kafka's tale of that name.

If we approach closer to real cases, we find that frameworks are never effectively total, and the user of one can always wield at the same time at least part of the armoury of the other.

Even Feyerabend implicitly admits this; as I have pointed out, in expounding Galileo's method of persuasion (by "propagandist machinations"), he shows that in order to get his auditor's to change their "natural interpretation", Galileo must first obtain their agreement - in the

tower argument, agreement on the not-always-obviousness of motion. This is already to have them change their framework, albeit very slightly. But of just such slight changes, sufficiently accumulated, large changes eventually result : for they make way for further changes in a direction already present - not in some platonic realm, but in the structure of Galileo's own framework, which is clearly the richer here - for it includes by correcting it the Aristotelian one (just as in the Newton/Galileo case above in 1.3). Motion, Galileo would allow, is often obvious, but not always, because ... and here follows his whole argument. (And it is an argument, Feyerabend notwithstanding.)

The framework of Galileo's auditors, I said, was changed : clearly I must regard frameworks as very mutable to say such a thing. That I do, and why, will be the clearer below when I discuss changes such as the foregoing in terms of the introduction of referential rules.

In short, some frameworks (partial sense) may be disjoint and so in a suitable sense genuinely different; but frameworks involved by a succession pair of theories cannot be : for at the least, the successor theory has the resources to discuss what it succeeds and its failings : and this requires semantic overlap - we cannot speak of what is wrong with mediaeval explanations of disease and madness without denying some of their ontology. Thus the premise (P7), on the partial interpretation of "framework" is true only for the weak sense of "genuinely different" (i.e. non-identical), not for the strong sense (disjoint) - which is to say, in precisely such case as makes the inference required (to (C5)) invalid.

It is doubtless points such as these which lie behind what Kuhn and Feyerabend say about incommensurability in their more guarded moments. For example, Feyerabend says "only suitably interpreted" are (some) rival theories incomparable.

What this means is that "A realist ... will use (these theoretical terms) to give meaning to observation sentences or else to replace their customary interpretation" ...³²

This sounds dangerously like a claim to do what the Tarskian theorem rules out (cf. p.23 above). Feyerabend's argumentation here is revealing. First he says that the objection against this that "theoretical terms have no independent interpretation" rests on the idea that

"new and abstract languages cannot be introduced in a direct way ..."

which he claims is refuted by the way children learn their first, and anthropologists a primarily unknown, language.

In neither of these cases is the language abstract (especially the parts learned first); and for the child the language is not a "new" one in the relevant sense. No doubt new and abstract concepts may be introduced in a direct way, but in the nature of the case they are introduced into a framework, thereby changing it. They are not introduced as a framework replacing the old one - as Feyerabend's own example (of the tower argument, mentioned above) clearly evidences.

1.7 I have left several loose ends in this chapter which dangle from the idea of rules governing discourse - in the argument of 1.2, in which a necessity for rules is deduced from rationality; in discussing Kuhn's claim that the use of concepts like simplicity in theory-choice is not rule-governed in 1.3; and the statement of the argument from frameworks in 1.4. The notion of rule is also required for the notion of framework in 1.5. Finally in 1.6, the argument that Galileo changed very slightly his auditors' frameworks was left with a promise to expand on it by discussing the introduction of referential rules. So the next chapter begins with a general discussion of rules.

CHAPTER TWO

RATIONALITY AND CONCEPTUAL REVISION

2.0 Preamble

Having a concept is not the same as having a reason, generally speaking, but it seems that having concepts is the same as having reasons in the sense that only a rational entity can have concepts, and only an entity which has concepts can have reasons. (This is not a paradox.) If only we could say just what concepts are, we could go on to characterize having them, and end up with something useful about rationality. This, what concepts are, is a notoriously unsolved problem of course, and we really can only proceed the other way - to try to characterize "having concepts" with our bare hands, so to speak - hoping this will throw some light on what concepts are. Obviously we shall get a much weaker light on "having reasons" this way, but it may still be illuminating. The most enlightening mode of talk about having concepts recently introduced is talk about rules of language.

As a matter of epistemological caution, we must first say what we can about having concepts in cases where they are definitely had, before we can hope to deal with more doubtful cases. Hence the focus of attention is the evidence of having concepts - to wit linguistic deployment of them. Later in this chapter, I shall say something about the dangers of taking this too far - having a concept cannot be cavalierly identified with "using a word". (At the very least, the latter phrase must assume as subtle a sense as that of concept if such a claim is to be accepted).

Thus, the evidence of having reasons is giving them : and this involves the exercise of concepts. It has recently been claimed that

the exercise of a concept is just following a rule : indeed several writers have suggested that rules provide the link between concepts and reasons in precisely this way : the rule which governs the exercise of a concept provides the reason for its use. This has been argued explicitly, for example, by Kovesi¹, and more concretely by Bennett².

The danger with this line of talk, I shall argue, is that to regard concepts as individuated by rules leads naturally to the notion that concepts do not change. Allied with the similar notion that logic does not change, this in turn leads easily to the conviction that reason does not change. This is precisely what I deny : reason is an evolutionary product, hence an evolutionary process, so intrinsically mutable. Part 2 of this essay is a positive development of this view.

In this and the next chapter I wish to argue against the two main props of the common prejudice that reason is immutable - in this chapter, the fixity of concepts; in the next, the fixity of logic.

The main point of this chapter is this : all concepts are mutable. It is a consequence of this that there cannot be ineluctable concepts - concepts that any rational being must have. This gives further point to my articulating the Urmutter fallacy - not only is that inference unsound, but its conclusion is false, at least in its 'concept' version. In point of fact the three versions cannot be separated, and the argument in chapter 3, that logic is reviseable, also supports my claim in this one - if all rules are reviseable (in Quine's sense),

1. Kovesi, (MN), chapter 7 passim.

2. Bennett, (R), passim.

as I argue in Chapter 3, so must be concepts since they all involve rules. *Pari passu*, the argument of this chapter supports the claim of the next : for if all concepts are revisable, so must be in particular logical concepts.

2.1 Rules of discourse

Talk of rules was greatly encouraged by Wittgenstein in the English-speaking philosophical world, but many philosophers who use the term are somewhat coy about just what it means. For example, R.M. Hare's Reason and Freedom³ leans on the notion considerably, yet he declares :

"It suffices for our present purpose to say that by 'rules'

I do not mean very simple general rules which can be formulated in words ... but, rather, that consistency of practice in the use of an expression which is the condition of its intelligibility"

This laudable recognition of the complexity of linguistic practice need not be tied to such vagueness, however. Hare goes on to recognise that the rules "determining meaning" need not be all of the same kind. I propose that indeed rules of various different kinds determine the meaning of any particular word - and, usually, lots of them. This is why a 'simple, general' formulation may not be available, I suggest - but, I would like to further suggest, each individual rule is

(a) general

and (b) normative

In fact, I propose that a rule is, precisely, a normative, general *type of* inscription (the latter term I take to include 'aural' inscriptions,

3. Hare, (RF), p.7

that is utterances). A rule is a linguistic item, but of course what it rules need not be language. The favourite examples of rules in philosophy are rules of games. (The psychology of this fact is interesting but irrelevant here.) Two rules are "the same" if they divide up possible cases the same way into compliants and non-compliants.

(Two further points about Hare : it is very clear from the text that the two marks of a rule I have given are what he wants with them - to provide his 'universalizability' and 'prescriptiveness' for moral judgments. A second point in passing : in discussing different kinds of term, he points out that one cannot misuse the word 'it' by applying it to an object to which it may not be applied (contrast red). Then he says that "Fortunately it is irrelevant to our purpose to inquire into the very difficult question of what would constitute a misuse of the word 'it'".

Hence are some examples of misuse of "it":

"Mother can't come, it has a cold."

"Cows it mammals."

"I have seven apples, but it is rather green."

Of course there are lots of other ways to misuse this word - my point is that each way indicates a different rule which is being broken - in these cases, the rule that applications of 'it' to humans are inappropriate, rules to the effect that 'it' can't act as verb, the rule that its reference is to an individual.)

It is quite common to recognise that the rules governing linguistic behaviour are of more than one kind. Particularly common is

the distinction drawn for example by Searle,⁴ Whitely,⁵ et. al., between syntax and semantics (under other names, usually).

Syntactical rules proscribe certain forms of utterance as not counting as linguistic units at all, in the way that the rules of chess rule out a diagonal knight's move as not part of the game (example : a sentence without a verb). Semantical rules constrain the sense which can be conveyed by the different linguistic units - they are the rules determining which objects are referred to by which nouns, and so on.

There is a third kind of constraint on what may be said, however, which is sometimes but not usually joined to these two. Sometimes they are referred to as pragmatic, sometimes as dialectical. None of the terms I know suffices because some of these rules are logical. A simple example is the rule against contradicting oneself, but not others. This rule would not commonly be called logical, but perhaps dialectical.

A different distinction which is of some importance is between explicit and implicit rules. In each of the three categories of syntactic, semantic and pragmatic, there may be examples of both. The syntactic rule of English, that sentences should not end with prepositions is fairly explicit, that is, many well-educated speakers could quote it as a reason for the way they speak. Yet other speakers can follow it, and also express unease about infractions of it, without regarding it as a rule (they were never taught grammar explicitly).

4. Searle, (SA), p.33 et. seq.

5. Whiteley, (RL).

In the case of semantic rules this is particularly obvious in the qualified credence which people give to generalisations involving some term - in so far as they are accepted, they consequently and subsequently exert normative force (the term must be used as if the accepted sentence is true); yet there is nothing wrong with changing one's mind without giving up the term, if the generalisation in question conflicts with others, or even with unformulated constraints on use - this latter is very common in the case of terms being given new explicit rules via scientific theorizing.

Similarly for dialectical rules, which are almost all implicit (very implicit!). Various common charges of a rhetorical kind ("Irrelevant!", "Leading question!" etc.) are perhaps closest to being explicit for this class of rules. Some people can cite what would be usually called logical rules, such as modus ponens. Beyond the lower predicate calculus, the lack of agreement about which logic is right (for example, among modal systems) shows how quickly the explicitness tails off.

(This discussion will become more concrete in the next section.)

I do not intend to argue here for, but just accept, the idea that someone can follow a rule, indeed accept a rule, without knowing it - there is no need for rules that are accepted to be explicit. But I do think that if a rule is accepted it can in principle be made explicit - else what justifies saying "a rule" is accepted? This viewpoint will be extended in ch. 3 in talking of the logic someone accepts. I do not claim that all important rules or logic that are generally accepted have yet been made explicit.

2.2 Bennett's Thesis

In this section, I give detailed consideration to the attempt by Jonathan Bennett⁶ to analyse the notion of reason, firstly because it is one of the few such explicit attempts, secondly because it focusses on the notion of following rules, and thirdly because it results in just the static conception of reason which I reject.

Bennett's thesis is this : it is necessary and sufficient for a community's being rational, that its members have the abilities to make general and "dated" statements. Just what a "dated" statement is, is best explained by sketching the way he arrives at this thesis. He described five (hypothetical) communities of bees exhibiting successively more sophisticated behaviour; only the last of these, he claims, must be recognized as fully rational, the other four approaching successively closer to rationality, but falling short.

The first community consists of honey-bees as described by Karl von Frisch. A bee which perceives food will, on returning to the hive, perform a dance which codes its location - the axis of the dance-pattern indicates the direction of the food, the frequency of the pattern its distance, and so on. Thus the relation between dance and situation is one of concomitant variation - the Tractatus model of language in fact. Bennett, however, denies that this is linguistic behaviour since it is regular but not rule-guided.

For rule-guided behaviour, reports breaking a rule must be recognized by the bees. So the second (imaginary) community have another kind of dance, called a denial, which a bee performs if a report (i.e. dance

6. Bennett, (R).

of the first kind) which it observes, conflicts with its perceptions. However, a report and its denial do not constitute a genuine conflict, since food may disappear in the interim between observation, report and checking by a second bee. To deal with this, for the bees in the third community a report 'says that' there was recently food-evidence (which may be food) at a certain place. A denial can be construed as the negation of such a report without irrelevance or miracle being implied.

This community, Bennett rather mysteriously allows (on p.45) "have reasons in a minimal sense", for there is for them no simple correlation between a bee's immediate environment and the dance it performs - "being evidence for ..." is a complex relation between dance and situation. Yet these bees are not rational, since the next community, though more sophisticated, are not. The latter may report in both past and present tense, about food or danger, and can give reasons : that is, append to a report, a further report of the circumstances which provide evidence for the basic report. (Example : "Was food at S; flowering daffodils").

The reason why even these bees are not rational is that these given reasons are only reasons if a denial can deny the evidential link - that is, if the copula is really stronger than the semi-colon - something like 'because'. Yet these bees cannot distinguish such a putative denial from one which simply denies one of the two constituent reports. So a further kind of dance is required, and furnished for the fifth community, which is called an R-denial to do just this job. In effect it states "that 'because' is no good". To eliminate the possibility that this dance does not state this, but is merely a synonymous variant of the standard denial, a bee so

challenged may demand, and be danced, a report of a counterexample to the rule implicit in the reason-giving dance (e.g. "At S at T, there were flowering daffodils, but no food" - a counterexample to the rule "There is always food where there are flowering daffodils" which is being invoked if the copula is really 'because'.)

Now although it is obvious that the bees of community two are not rational, for they simply sometimes contradict each other and nothing more happens, what is not so obvious is how a bee from the fifth community whose reason-giving dance is R-denied, is any better off. What makes the difference, Bennett claims, is that a bee so denied-to, can decide rationally whether to maintain his report by continuing to use the implicit generalization, or to give it up by accepting the reported counterexample.

This can be done because the denying bee can be asked for, and can give, more detail about his alleged counterexample. What this means I discuss below.

To sum up : in the third hive, there is linguistic behaviour, for it is rule-guided since genuine denials occur. Reasons are had by bees, though not given. In the fourth hive this lack is remedied through the provision of (implicitly) general statements. But though "reasons" can now be given, that they are reasons is only certain if the implicit general statement can itself be contested - for which role the bees of the fifth hive require "dated" statements. That is, not only can they report a counterexample (there is nothing new in this - it is simply a report), but also add detail to it and (when in the other position) evaluate such a story. They must be able to make and assess dated statements, as Bennett allows (on p.72).

But he goes on ... "I do not think that (this) would introduce radically new elements."

This is a puzzling claim : for on the face of it we are in this dilemma - either, as Bennett claims, the ability to assess dated statements adds nothing new, in which case the alleged distinction between linguistic and rational behaviour seems to collapse. For if this ability (assessing over and above making dated statements) requires "no radically new elements" then these bees are the same in capacity as those of hive four. Or this is not so, in which case we are entitled to be told in what this extra ability consists. Either way, Bennett's thesis seems to founder.

The crux of the matter is the contention that a bee which dances a reason-giving dance and is R-denied, can probe the alleged counter-example and therefore decide rationally what to do. Let us see if there is any reason to accept this inference, granting a certain latitude to what a probe might be, to allow for Bennett's admission that they "need a much less crude set of principles of the form' ... is a reason for ...' "than they have. (p.72)

The situation is this. B1 makes the claim

(1) Food at S1 because circumstance C at S2.

(It does this by dancing a particular kind of food-dance).

B2 R-denies this, B1 requests an explanation for his rudeness and

B2 dances a counterexample:

(2) No food at S1*, but circumstance C at S2*, and S1* was related to S2* as S1 was to S2.

(the third clause is necessary to avoid jokes like "daffodils are no reason for expecting food - why therewere some in Coleman's garden last spring but there's no food there now".)

Now B1 cannot just accept this report (2) and at once give up his 2-part claim ((a) Food at S1, (b) because ...), that would be irrational (unless perhaps he had only just floated this hypothesis - but Bennett does not explicitly consider this element of the process. It is clearly going to be required unless, magically, the bees sprang into existence knowing all about food and danger. I will return to this problem.)

Assuming that the implicit rule

(3) Whenever C at S2*, food at S1* is entrenched, B1 must ask for more detail; perhaps he does this by repeating (1) a few times "dogmatically". Now what does B2 do?

Obviously, his response must be to make further reports, perhaps reason-giving ones. I suggest that there are only two ways these can be related to (1), (2) and (3). On the one hand there can be a breakdown : the pair (F1*, C2*) is buttressed by a series of at least two pairs (F11*, C21*), (F12*, C22*) ... That is, B2 gives a series of reason-giving dances like (2) 1 : sub-food-fact F11 at S1* but sub-circumstance C21 at S2* and so on, where "F11 and F12 ..." is a more detailed version of "Food Report F1". (For example, F11 might be 'no smell of pollen'; F12 'no sugar solution' etc.)

Secondly, (2) may be founded by a dance of the kind

(4) : Circumstance C at S1* because circumstance CC at S1* (for example "it was a flowering daffodil because it had a yellow trumpet".)

Any good probe will create an evidential net involving both these moves. I do not see that any others are available, however.

If this is so, we must now enquire how B1 is better off after he has elicited some such net of detail. Originally he was faced with a choice between rejecting (a) the rule he was relying on in his original reason-giving dance, or (b) B2's report F1*, or (c) B2's report C2*. After the probe, B2 has made some more reports (at least one) and invoked at least one more rule (one covering (F11*, C21*), another for (F12*, C22*) etc., or one covering (C,CC). There are various possibilities :-

(i) B1 may "know" from his own experience that one of B2's further reports is mistaken - but in general this cannot be so, for in general B1 was not at S1* or S2*.

(ii) B2 may invoke a rule not generally accepted in the community, in which case B1 may either challenge it, leading to a further probe, or simply use this as an excuse to disregard B2's whole story; but if it is envisaged that such deviant rules are not strictly anomalous, such a response will simply (eventually) split the community into separate subcommunities - a kind of apian conceptual relativism which is also clearly irrational;

(iii) but in the more standard case where neither of the points (i) or (ii) apply, there seems to be no way in which B1 is better off - rather he has a wider choice of irrational options, that's all. Moreover, any of the rules and reports which are elicited by his probe should in general be at least as safe as those in B2's original counterexample, since, generally, only the latter directly conflict with B1's own claims. Since only conclusive reasons are allowed to the bees, (pp 59-60), it is just not true that in hive five they have

the notion of an implausible story as Bennett claims they do on p.73, for that requires, clearly, the relative weighing of the various reasons given in a probe.

A parallel, and a suggestion

We may see that this is the point at which Bennett's scenario is deficient if we step outside his apian framework and briefly consider the scientific parallel he has in mind. The problem facing B1 is precisely that posed as the "Quine-Duhem thesis" in the literature of the philosophy of science : given a set of laws and a set of observations, which cannot all jointly be correct, there is no deductive procedure for deciding what must be rejected. I think that the answer to this 'problem' is that such situations are in fact controlled by rational but non-logical criteria, such as simplicity, scope and so on. These essentially involve the relative entrenchment of various "laws" and "facts".

Returning to the apian fable, it is clear that precisely such criteria have been summarily excluded by Bennett's furnishing the bees with a notion of conclusive reason, but not the weaker notion of fairly strong evidence. (p.60). He claims that this is acceptable since the latter notion is logically anterior to the former. But is it? His only argument for this claim is that

"one could not have ... the concept of that which provides some reason for saying that p, unless one had ... the concept of that which makes it downright wrong to deny that p" (p.60)

If this is so, how is the latter concept acquired? (But perhaps it

isn't.) In any case, if the claim is true, then surely the introduction of "reasons in a minimal sense" in hive three (reports and denials of food-evidence) is illusory. For if denials are to be conclusive, then a given circumstance just is or isn't evidence of food. In which case the correlation between circumstances and dances covered by the rule which guides the bees is only complex in the uninteresting sense of complicated : the suggestion on which Bennett leans that these bees assess evidence is unwarranted. For in this case the rule could presumably be fully formulated in total accuracy : in which case one would be justified in looking for just such a straight forward stimulus/response account as is alleged to be ruled out.

For what force does "assess" have in this situation which goes beyond the ability to respond of the very first community of bees? Precisely what is lacking, is any notion of weighing alternative responses, which is essential to the notion of assessment, which is cognate (conceptually) with judgment.

Nevertheless, that this is what Bennett has in mind can, I believe, be shown by considering the things he says about the various rules which enter the apian story.

In pursuing his fiction, Bennett takes it as hardly being in need of argument, that to be rational the bees will need to be rule-guided. He does argue that to be rational a creature must manifest behaviours "appropriate" to what is not both present and particular in its environment.

Consequently,

"The special power of language in these matters derives from the fact that linguistic behaviour is behaviour which obeys rules correlating performance with empirical states-of-affairs."(p.88)

Still, Bennett's thesis is not that rationality is linguistic capacity per se, so some conditions on the rules would appear needed to back his claim. In fact three different kinds of rules are involved in the apian fable, though he does not draw attention to this fact.

The first kind of rule, which I will call rule type 1, is that mentioned in the quote above - the rule which guides, for example, the bees in the third community : a generalization correlating dances and states-of-affairs (a generalization such as merely described the first hive), which has assumed normative force for the hive three bees and guides them. Bennett, however, points out that we are not entitled to say that these bees have rules (p.57); still, if our description is correct they do. (His methodological assumption that no meaning can be attached to this last caveat will turn out to have significant consequences).

His reasoning here is very odd: the problem he claims, descends from the need to give 'rule' "an unassailable place in the description of the bees' behaviour", which he attempted by providing denials. This having failed, he now proposes to solve the offspring problem with the same stratagem-introduce R-denials; this certainly sounds unpropitious.

However, the effect of the last move is to ensure that the bees must be granted to be following the rule (which we thought the hive 3 bees followed) because they can quote it. The question is, is it the same

rule? There seem to be two, not necessarily conflicting, accounts of the relation between these two rules.

On the one hand, we may say that it is the same rule, which was implicit for the hive three bees but is now explicit for the hive five bees. The one rule has been "brought before the mind".

Alternatively, we might deny that a single rule is involved, for the following reason : for the hive three bees, the rule we guess to be guiding them is a generalization over dance episode - state-of-affairs pairings. The rule which the hive five bees use to justify their food-reports, however, is a generalization over dance-episode - dance-episode correlations. In other words, the hive 3 rule relates the concept 'food' to its instances, but the hive 5 rule relates the concept 'food' to the concept 'flower'. Thus there is a contrast between a semantic rule for hive 3 and a syntactic one for hive 5.

Bennett has some discussion of matters bearing on this question in Kant's Analytic,⁶ in a passage closely connected with the whole enterprise of Rationality. He dismisses the possibility of semantical rules, there called referential. Before discussing that, however, I wish to point out a further kind of rule undeniably claimed to be guiding the bees.

On p.78, Bennett refers to the "adjudicative rule" which a bee in a quandary (i.e. faced by an R-denial plus probe confronting its reason-giving dance) must follow in order to resolve the conflict.

6. Bennett, (KA)

He dismisses various possible rules, such as that R-deniers are always right, or always wrong unless trusted etc., on the grounds that such rules make the bees' community irrationally based. Clearly, the specific kind of rule required here is crucial. What kind is it? Bennett gives no positive account, implying that none is needed (no "radically new elements" p.72), yet it is at least prima facie evident that this rule cannot be of either of the kinds previously given the bees.

This rule does not correlate dances and states-of-affairs, nor dances and dances. Rather, it must be a generalization covering pairings of linguistic interchanges on the one hand (confrontations) and changes to the bees' stock of accepted rules and reports on the other (dispositions to dance, rather than dances.) It looks very like what I have called a pragmatic or dialectical rule.

2.3 Applying Concepts

Bennett discusses explicitly the role of rules in the application of concepts in Kant's Analytic, §35-37 particularly.

The general account of concepts which I dispute regards concepts as individuated by rules. Bennett's early concern with altering the bees' behaviour is concerned to make it rule-governed. One way of putting his view is this : a concept is an ability to use a word properly; proper uses of a word constitute a rule which states all the circumstances in which it is properly used. The rule may be a disjunction of simpler rules, but none of those may be a referential

rule, since a rule relates the applications of the concept to each other, it does not relate the concept to an (or even every) application. There may be difficulty in specifying precisely what the rule is for any particular concept, but that is neither here nor there : in principle it can be done.

It is clear that referential (semantic) rules, were they possible, would cast doubt on the individuation-by-rules thesis, since any instance of a concept must be an instance of other concepts too. Now I claim that ostensible counterexamples, that is, new putative instances, are frequently the occasion for the possibly temporary but nevertheless necessary introduction of referential rules. A simple example is the well-known case of people who had the concept 'swan', and were suddenly confronted with (for the first time) black swans. I think we must say that however else you describe what happened, a rule was introduced to the effect that "black swans as found in West Australia, i.e. certain particular birds and any others like them there may be, are to count as swans".

In Kant's Analytic discussing the schematism of the categories, Bennett rejects the question Kant is addressing, viz., "How are concepts applied to their instances?" as senseless, and rejects, as an equivalent error, that there can be referential rules as well as non-referential ones (the contrast is Körner's, Kant p.71). The former allegedly relate concepts to their instances, the latter to other concepts. The argument against this possibility is that an absurd infinite regress follows from it, as it allegedly does from the proposition which Bennett treats as equivalent, namely that there could be "a technique of concept-application as such". He also employs directly against the

notion of a technique of concept-application, the argument that it would be redundant, since "having a concept involves being able to apply it." However, these arguments do not really stand up to examination.

Bennett argues (p.145,KA) that there cannot be referential rules.

For suppose we have a putative referential rule :

(*) You may apply concept C to a thing just as long as the thing is ...
 "ending with some description." (say D)

But, he points out, the phrase "just as long as the thing is ..." is tantamount to "if you are also prepared to apply 'D' to it".

Which looks like a non-referential rule relating two concepts, C and D.

He hedges this rather abrupt volte-face with the admission that it might still be referential for some person at some time, depending on what he knew, but not referential tout court. For the present I wish simply to note that this implies that a non - referential rule is, then, non-referential tout court.

Here are three objections to Bennett's argument.

Firstly, the way it is formulated in the second person ("You") insinuates without actually claiming, that a rule must be so formulable, and this is equivalent to the assumption that every rule must be explicit : a rule which guides me, which I cannot state, can hardly be so formulated, for this expression entails my (overt) awareness of it, even if intermittent.

Secondly, the use of 'only if' in Bennett's formulation implies that a rule must be prohibitive. But this is not so, rules may also be concessive : a rule might be of the form

(**) You may apply concept C if

After all, the rule as Bennett gives it is perfectly consistent with a set of other possible rules, excluding other possibilities, such that no circumstances were not excluded one way or the other. But the fact is that workable discourse contains presumptions in favour of its happening - no discourse would occur if the only rules controlling it ruled out some possibilities and nothing ruled others in.

Thirdly, it is not obvious why the rule must end, as Bennett suggests, "with some description" - why not with a name, perhaps? Why can there not be ostensive rules? For example, presumably there are rules governing the use of the term "Pope". Surely amongst them must be one to the effect that Peter was the first? (This would still be general, applying to all uses of the compound concept "first Pope" perhaps.) Further examples of needed ostensive rules will be given shortly.

In the same passage, Bennett makes it clear that he regards the preceding argument as equivalent to the thesis that "there could not be a technique of concept-application as such ... since the implementing of a technique requires the application of concepts" (pp.144/5). Thus, it is incoherent to suppose that each concept-application presupposes an application of some particular concept, C, since this would give rise to an infinite regress.

I admit that it is absurd to suppose that there must be an infinite hierarchy of identical concept-application concealed within the one of interest, like some Borgesian onion. (The line of argument is clearly related to Rylean objections to the homunculus theory of the mind.) However, this is a fairly weak admission once the point is contested that identical concept-applications are presupposed simultaneous

with the basic one. It is true that applying C does presuppose applying C and so does that ... But the presupposition is temporal, and since the C is not identical from application to application, there is no inevitably infinite regress, just a very long one going back to childish babbling. Concepts are formed; they change in the process of formation and it is arbitrary to suppose that at some point a line can be drawn, the concept is there and will change no more. But just this is what the infinite-regress argument presupposes.

Bennett has another argument against there being referential rules - not only would they lead to vicious regress, they are redundant anyway. It does not make sense to suppose that I have a concept but cannot apply it to its instances.

"I could not possess a concept yet be unable to apply it ..."

(p.146) "(If I can) state many general truths about dogs, such as that they are mammals, never laugh, have legs, etc. If I can do this and yet - although not sensorily disabled - apply the word 'dog' to particular birds, humans, porpoises etc., and often apply 'not a dog' to particular dogs, you must conclude that I do not understand 'mammals', 'laugh', 'legs' etc."

Must one? Could not I have learned all this from a traveller's tale, dogs never being seen in my country? Is it not equally plausible to lay the blame on my not being governed by any referential rules, in short no one ever having pointed out a dog to me?

It does seem strange, nevertheless, to suggest that this situation could be sustained if explored (such situations generally aren't

explored in real life, as opposed to philosophy classrooms); but in any case, it is enough to point out that Bennett's claim is too strong, if it means that if I have a concept I must be able to apply it to all its instances. Otherwise no-one would ever learn, for example, that a Great Dane is a dog, not a bizarre pony - for someone who needed to be told wouldn't on this view have the concept 'dog' and so couldn't be told. In any case, if one considers concepts less homely than 'dog', the claim begins to lose both plausibility and indeed significance. Does anyone have the concept 'philosophy' or 'wickedness' on this view, except, perhaps, Plato and Satan respectively? Does anyone at all have the concept 'quark'? What about some system of concepts in a burgeoning area of research - say 'inflation', 'price-rise', 'seasonal effect' etc.?

Bennett's is clearly an exaggerated claim. Consequently, in general, a concept may be had while yet its possessor regards some putative instances as problematic. Which is why B1 may be in a real quandary rather than just pinning down a slip on someone's part. Particularly acute problems may be posed by a whole new kind of instance arising. An example of this is the black swan case referred to above.

The possible argument as to whether the explorer had one concept and then another or had only the one which changed brings me to Mr. G.J. Warnock, who in an article⁷ which Bennett very approvingly quotes in his discussion of concept-application and who has very similar ideas about concepts attacks what he regards as a "dangerous terminology", to wit the talk about people "having concepts".

7. Warnock, (CS)

Now this was way back in 1949, so perhaps people were more prone to be silly then. Still, I find it implausible when Warnock says that talk about "having concepts" lends itself to thinking of a concept as an object - his example of an object with which you might confuse a concept is "a gauge for measuring knitting needles" - and that such talk "suggests a variety of impossible questions". Now some of these questions are a little odd - for example "Where are my concepts now?" Others of them however seem to be perfectly possible, indeed sensible, and to have obvious answers. "We might even ask whether there are concepts which no one has ever had, or which do not apply to anything at all". (presumably 'even' indicates that these questions are more impossible than the previous one) But those questions seem to me to have the obvious answers "yes" and "yes". (Concepts which no one has ever had will include those that have not been invented yet, just as in 1900 no one had ever had the concept of "semi-conductors"; and concepts which apply to nothing at all might include "God", "unicorn" or "square circle" - depending on your prejudices.)

I have introduced this not simply to make fun of Warnock, however; in my opinion the facts that concepts can apply to nothing at all, or may not have been invented yet, are important : for if concepts were not inadequate (by having no instances, or in other ways), there would be no need for conceptual change; and if there could not be concepts no one had ever had, there would be no scope for it. This would be a pity since it happens, and the growth of knowledge is not just a cataloguing operation. So I think Warnock's fears are not only exaggerated but pernicious.

What he suggests is that we talk about being able to use a word, rather than "having a concept", in order to avoid these impossible questions. (I shall not stop to consider "impossible questions" that may arise from this substitution - for example, must one party in the argument about black swans be unable to use the word "swan"?) The effect of it on Kant's question as to how concept-application is possible is "to ask how I can use a word that I know how to use. And this is a silly question, brought on by illegitimately separating the application of concepts from having them." (This is one of the passages Bennett approvingly quotes.)

It is not obvious that it is a silly question; for example consider the black swan case again. But what I want to pursue is a shrewd observation of Warnock's about bicycles and monkeys. He asks whether it is fair to say of a monkey which can separate out all the white ones from a set of variously coloured blocks that do not otherwise differ, that it has not the concept 'whiteness' since it does not use the word 'white' (or any other)? He compares this with the case of my having had a bicycle which has had the wheels stolen - have I a bicylce or not? He suggests that 'we maintain that having concepts presupposes knowing a language - with the reservation that something more or less like having concepts may occur where no language is used.'

Now I want to argue that this suggestion - which is, in effect, one which Bennett adopts - is arbitrary. To show you why I say this, consider an example of Ryle's⁸ - namely the question : when does one say of someone that he "has the concept" of evenness (of numbers)?

8. Ryle, (TTAHC)

A child who can count is told that each number is odd or even; then that 2,4,6,8,10 are even, 1,3,5,7,9 are odd; next that the number on the left-hand page of a book is even (!); then that any number ending in 1,3,5,7 or 9 is odd; then that any number next to an odd number is even; finally, that any number which divides by 2 is even.

Ryle stops here and asks : when does he have the concept of 'even number'? He concludes "the question has he really got the concept or has he got the whole of the concept so-and-so? is like the question has he really learned the art or has he learned the whole art of skating?"

I think that this sort of case shows that even presupposing language, having a concept is not an all-or-nothing affair - which is why I call Warnock's suggestion arbitrary : there can still be more to having the concept of evenness - for example, what about fractional numbers?

What would need to be true about concepts for all of what Bennett says and implies to be true? First, that concepts be given (to new bees), fixed (every application is the same) and adequate (unimprovable, guaranteed instances and without problematic cases). For this to be the case, two other conditions would have to be fulfilled : Firstly, that the bees are omniscient about their little world (else any problematic case would raise the possibility that the rule was wrong seriously; but if the bees genuinely have the concept 'food' they must, on Bennett's account be able to apply it to its instances. Which means they must know what food is and what flowers are and the reports cannot be in jeopardy (except for accidents to the bees' senses etc.) - so in every case the rule must be the loser. But if this is the case, what is its point? If it can't sometimes be right, it has no function. And secondly that their world is Fregean - that is, if their concepts are going to be adequate, everything must

be definitely an instance of a given concept or not. Vagueness must be non-existent or at least fully eliminable. These last two conditions seem to me to require a world of limited complexity, much different from the real world.

In the next section I shall give some evidence, gathered from the real world, that concepts are not the way Bennett and Warnock would have us believe; fundamentally that they change, and as precondition of this that conceptual pathology should be a philosophical discipline at least rivalling in importance its more straight-laced neighbour, logical geography.

2.4 Conceptual Change

It is possible that at various stages in the development of his thought, one could have said that someone had the concept of 'number' at any of Ryle's stages, though this may be disputable. Consequently I now offer an example where the parallel statement is, I think, not in dispute.

It concerns the concept of poly hedron, and my remarks are based on Lakatos' brilliant series of articles "Proofs and Refutations"⁹

The concept of polyhedron

I think most people "have the concept" of polyhedron, in the sense that they dimly remember some stuff about solid geometry - the most obvious examples being cubes and pyramids. One of Kepler's obsessions was trying to fit certain regular polyhedra inside one another as a model of the solar system.

Let us say to start with that a polyhedron is a solid figure bounded by planes. Now the main features of these things are vertices, edges and faces. Each face is a polygon - that is a plane figure bounded by straight lines - the cube has squares.

Now polygons are, as their names indicate (pentagon, hexagon etc.) classified by a numerical prefix, where the numerical prefix refers indiscriminately to the number of edges, the number of angles and the number of vertices. The fact that the number of edges is the same as the number of vertices allows this convenient method of classification. Now a problem some people found interesting was the problem of similarly classifying polyhedra (it can't be as simple since there are obviously cases like a pyramid on a square base, as well as one on a pentagon and so on).

In 1750, the great mathematician Leonhart Euler came up with the formula

$$V - E + F = 2$$

and suggested that this could be of some use in the problem, which he was himself at that time working on.

Now this handy formula clearly covers the most obvious cases, for example :

cube :	$V=8, E=12, F=6$	$V-E+F = 2$
--------	------------------	-------------

triangular pyramid :	$V=4, E=6, F=4$	$V-E+F = 2$
----------------------	-----------------	-------------

square pyramid :	$V=5, E=8, F=5$	$V-E+F = 2$
------------------	-----------------	-------------

and so on.

Here is a simple "proof" of the conjecture that any polyhedron obeys this formula, i.e. the generalization,

$$\text{for any polyhedron, } V-E+F = 2$$

I do not aim at rigour, only to get the idea of the proof across. It is due to Cauchy.

Proof: There are three steps.

1. Imagine P made of rubber; remove a face;
stretch the rest out onto a plane without tearing anything.

Vertices become vertices, edges become edges and a face becomes the interior area of a polygon.
Removing a face reduces F by 1,
stretching doesn't change any other number.
2. Triangulate

(each line added increases E & F by 1, so $V - E + F$ remains the same)
3. Remove triangles one by one :

this doesn't change $V - E + F$

Since the last triangle has $V - E + F = 1$, and we discarded a face the original polyhedron had $V - E + F = 2$, QED

Now this is very cunning, but unfortunately, there are counterexamples.

Counterexample 1 : nested cubes

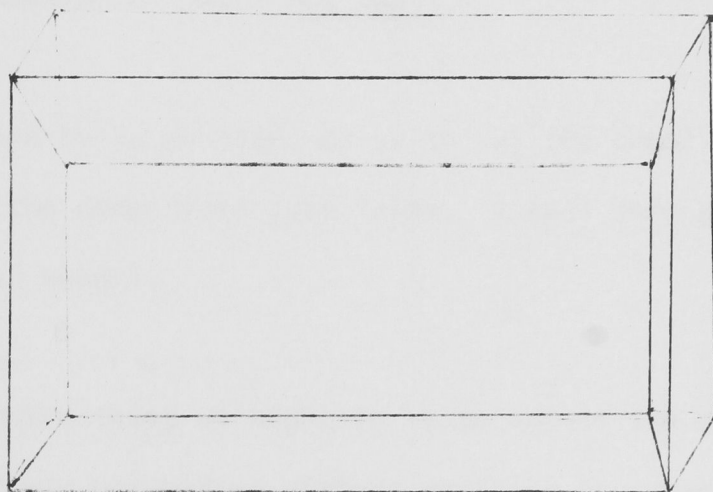
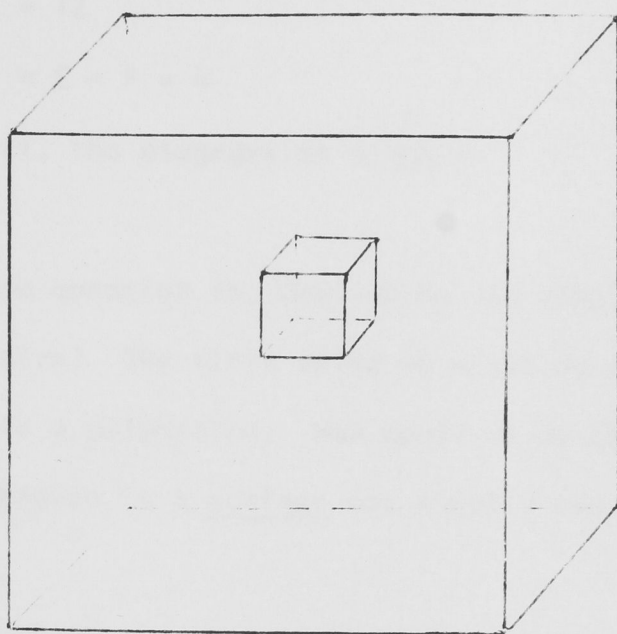
$$V = 16$$

$$E = 24$$

$$F = 12$$

$$V - E + F = 4$$

Counter example 1 : a cube with an internal cubical cavity.



Counter example 2 : a picture frame made from material of triangular cross-section.

Counterexample 2 : Picture - frame

$$V = 12$$

$$E = 20$$

$$F = 12$$

$$V = E + F = 4$$

(cf. the diagrams on p.57a)

Now, the question is, what do we say about this lamentable state of affairs? The first thing we might do is say that neither of those is a polyhedron. How could we do that? Well we might say a polyhedron is a surface not a solid (so CE1 is 2 polyhedra).

And we might say that if you put a plane through a polyhedron, you get one polygon - not two - so CE1 is not a polyhedron.

Lakatos calls this the method of monster-barring. (cf. declaring West Australian swans not swans.)

A second thing we might do is to say the proof is obviously no good, the conjecture just false. I call this giving up. (Talk no more of swans).

A third thing we might do is to adjust the conjecture, thus :

For any polyhedron without cavities or tunnels,

$$V - E + F = 2$$

This is called exception barring. (cf. "swans outside Australia are white")

A fourth thing we might do Lakatos calls lemma-incorporation.

To see what this means, consider why CE1 and 2 are counterexamples - the reason is that if you remove a face, you can't stretch onto a plane; so that is where the proof breaks down for these cases. So we could define a simple polyhedron as one which can thus be stretched when a face is removed, and put up the alternative improved conjecture:

For all simple polyhedra, $V - E + F = 2$.

This is not the end of the story by a long way but it is enough for my purposes. (In the swan case, the parallel move would presumably require that certain genetic-theoretic qualifications be built into a new generalization about swans' colours.)

What the polyhedron case shows

(Let me say that what the mathematicians did was take the fourth alternative) I think this example shows the following things :

- i) concepts can be inadequate - as that of polyhedron was when Euler put up his conjecture (and after);
- ii) concepts change, indeed develop;
- iii) referential rules are needed to develop concepts - the "monsters" must be declared to count as polyhedra;
- iv) a counterexample is more than a mere contradiction (hence the possibility of improving the conjecture);
- v) "having a concept", for some concepts at least, is a radically incomplete achievement;
- vi) there are more than two strategies for dealing with counterexamples.

This last point allows us to reenter the apian world.

Bearing on bees

Now, remember B1? I think we can make these considerations relevant to B1. Assume that bees are interested in geometry and pretend that B1's problem was to be faced with the generalization:

$$\text{For all polyhedra, } V - E + F = 2$$

and the counterexample 1.

What can B1 do? The way Bennett tells the story he may take either the first or the second of my alternatives - declare the CE a monster, saving the generalization; or accept the CE and drop the generalization. Now of course, Bennett does not rule out putting up new generalizations - else the bees would eventually lapse back to stage 3; but how do they do it? Perhaps on Bennett's account they could put up the first improved conjecture, though it is not obvious how they arrive at it.

It seems to me even less plausible that they could put up the second improved conjecture. But even if they could, my point is this : how would they decide which to put up? This is a question about how to apply concepts, and on Bennett's account consequently not a rational question. But I claim that the fourth alternative is more rational than the third. Why? Because exception - barring decreases content in a way that a lemma-incorporation does not - the second improved conjecture explains why the counter-examples fail to satisfy the formula; the first simply excludes them ad hoc.

CHAPTER THREE

RATIONALITY AND LOGICAL REVISION

3.0 Connecting Summary

The relation of this chapter to the previous two is as follows. The argument from frameworks relies, for any cogency it may have, on the idea that there is an essence to rationality, some necessary cognitive element which any rational being must have. The idea that certain concepts might play such a role was attacked in chapter two; in this chapter, the parallel thesis about logic is criticized. In terms of rules, in chapter two I pointed out how simple it is to change a framework by adding (or subtracting) a semantical rule. Later I shall give a little attention to the idea that changing beliefs (in particular, theories) changes the syntactic rules on discourse. In this chapter I want to argue for Quine's revision thesis - which for our purposes means that we are free to keep or reject any rule, even a logical rule (even the law of the excluded middle). As I pointed out in chapter two, this argument reinforces that of chapter two in that we would certainly be changing our logical concepts in revising logic. Indeed, I argue that the model of concept - change given in chapter 2 can be transferred to logical concepts without significant loss : although referential rules are no longer directly involved, nevertheless the basic pattern of minor concept change through rule revision remains.

The chapter may seem a ponderous application of Wittgenstein's "family resemblance" idea to the concept of logic. In fact, there is a good deal of truth in this remark, but only construed in the light of part 2, where the justification of that idea is argued for - to be blunt, families have resemblances for causal reasons which we can, and do, express in the theory of evolution. Similarly for other things with family resemblance.

The horse must pull the cart!

Quine's revision thesis is

(R) "Any statement can be held true come what may ... no statement is immune to revision."¹

If we adjoin to this claim the further thesis that

(P) Any rule which guides discourse can be stated as a proposition

then we may derive the conclusion that on this view any rule, even a logical rule, may be revised. Indeed Quine explicitly admits the possibility in the passage referred to. But what does this possibility amount to? One might be content with the conclusion that some logical rules can be revised; or one might wish to take Quine at face-value and argue that any logical rule may be revised. I shall argue that both these claims are sound, but that there is at least one kind of rule which, though revisable, may be abandoned only at the price of eliminating logic, and hence the continuation of scientific activity. Not that this implies that such a course is therefore irrational; on certain views of spiritual development it would be supremely rational if possible.

(It does have the peculiar property, if it occurs, of being the last exercise of rationality for the person concerned).

In the next section, I shall argue that there are clear examples of logical rules whose revision is perfectly intelligible; my examples will be quite standard ones, appealing particularly to intuitionism and

1. Quine, (TDE), p.43

quantum mechanics. In the rest of the chapter I shall try to show what is involved in logical revision by answering four objections to the idea : that revision is only pretended (as, say, for intuitionists "abandoning" the law of excluded middle); that revision is only notational, hence trivial; that it must be irrational; and that it can only be partial. It will be apparent that these four objections make successively greater implicit concessions to the idea of revision to logic.

3.1 Examples of revisions of logic : Intuitionism, quantum physics, modal logic, life-predicates.

I shall mean by "classical logic", L, the standard proposition calculus and the lower predicate calculus, as given in standard textbooks such as Hughes & Londey. I assume that we accept classical logic at present. Any logical calculus can be formulated in various equivalent ways. Normally one has some axioms and some rules of inference - for example a common version of propositional calculus has 5 axioms and 2 rules.² It is quite possible to 'trade off' axioms for rules; indeed axioms are not essential, although rules are. The revisions to logic which have commonly been canvassed are usually given as changes to the set of axioms, but this is an inessential point. A revision to logic consists in accepting one calculus in place of another; thus a glance at the logical literature reveals a plethora of putatively possible revisions. Some of course may be trivial.

My first two examples of revisions of logic are revisions to the propositional calculus. The first example is that of intuitionism :

2. e.g. as in Hughes and Londey (EFL) ch. 15

standardly, one regards adopting intuitionism as abandoning the law of the excluded middle. This is equivalent to dropping the elimination rule for negation

$$(DNE) \quad \frac{\sim \sim P}{P} \quad \text{(That is, "from } \sim \sim p \text{ you may infer } p \text{.")}$$

So this suggestion amounts to revising the rule DNE, that is, giving it up. The reason for wanting to make this revision is that it enables one to avoid the set theoretic antinomies.

A second possible revision to the propositional calculus is :

Abandon the distribution of 'and' over 'or', that is the law

$$p \cdot (q \vee r) \rightarrow [(p \cdot q) \vee (p \cdot r)]$$

This has been suggested by Putnam³ following Finkelstein⁴ as the best way to deal with the well known epistemological problems posed by the 'Copenhagen interpretation' of quantum theory.

(A 'rule' version of this revision can easily be given.) That this is a different revision to the first one is clear from the fact that $(p \vee \sim p)$ is part of Putnam's remaining calculus.

The logic obtained by taking Putnam's suggestion is not very well-known, unlike the intuitionist logic I. Clearly it does not exhaust the range of alternatives; indeed the extent of that range is not quite clear - how many (different) axiomatisations of L are there? Under what circumstances is the result of dropping a particular axiom from one axiomatisation the same logic as that obtained by dropping some (other) axiom from another axiomatisation? These questions, though interesting, cannot be pursued here.

3. Putnam (ILE)

4. Finkelstein (MSL)

A different kind of revision is demonstrated by the search for extensions of classical logic, as in debate about what is the 'right' modal logic - S4? system B? - here again the assumption that one of them is right if only we can find out which seems a prejudice unjustified by the weakness of our conflicting intuitions and the expanding material on which we can at any time employ our presently favoured system. This case demonstrates the danger of thinking that we accept a determinate logic even aside from the issue whether if we do, it is the 'right' one.

Perhaps for a given person at a given time, there is some determinable modal logic accepted, though even this seems doubtful; that there is no community-wide agreement is clear enough from the extent of debate about modal arguments and their propriety (for example, Anselm's Ontological argument for the existence of god).

This example exhibits a dialectical trade off between rules and cases whose frenzy is only an exaggeration of the true situation for the quieter parts of logic. Why should we think that these less well settled parts of logic are therefore less central than the lower predicate calculus?

Another example of a revision to classical logic, this time to the quantification theory, has been floated as a possibility by Dummett in connection with the problems raised by vagueness.⁵ (Other remedies have also been suggested). This is to give up the rule sometimes known as "mathematical induction". I shall myself be arguing *in part two that* this is required, on the ground that reasoners must be able to reason about reasoners and that life - predicates like "reasoner" are necessarily vague; so I say no more about it here.

5. Dummett, (WP)

3.2 First objection : revision is pretended.

The strongest kind of objection to the proposal that logic be revised is to say it is impossible : revision can be, at most, pretended. For instance, it may be said that intuitionists though genuinely abjuring DNE in mathematics have not really abandoned DNE; rather, for the strictly limited purposes of mathematics, the rule is suspended. However, it goes on, in order to use a deviant logic requires the use of classical logic in order to pick and choose when, and when not, it is to be preferred. Thus, a deviant logic can only be a person's secondary logic; the primary logic must be classical. I have three points to make here.

Firstly, note that the rule is not merely suspended : for intuitionists, a statement like

"There is a sequence of 100 7's in the decimal expansion of π , or else there is not"

is not true; but in classical logic it is true. So the difference does not consist merely in not drawing possible conclusions. This is an example of the interaction of logical and semantical rules.

Secondly, it would seem that the argument depends on two assumptions : that some spheres of rational activity necessarily require classical logic - unlike mathematical activity which seems to be possible with L or I; and that some such activities are unavoidable for a rational being. But the first assumption goes too far - for the second revision to classical logic, suggested by quantum mechanics, is one where $(p \vee \sim p)$ is retained, yet is not classical logic. Yet physical theory is just the kind of sphere of rational activity the

objection under consideration has in mind : for example, Körner⁶ contrasts the suitability of L and I for "factual" and "constructive" thinking as a kind of formal heir to the theoretical/practical reason distinction. An example like Putnam's suggestion makes clear the oversimplification involved. Moreover, the second assumption seems unjustified; is it inconceivable that a group of intuitionists should arrange to have their physical lives so automated that they can devote the whole of their time to constructive mathematics? Some mathematicians seem to achieve something like this for a time without cybernation!

An obvious riposte to this is that exercise of classical logic will be involved in the life support maintenance systems, quite likely on the part of other people : but this is irrelevant to the question whether some rational beings can dispense with L.

Thirdly, I would like to question the idea that some logic must be primary over others in a person's categorial framework. This kind of notion seems to identify a logic too closely with a calculus : I and L are distinct as calculi because they are not identical. But for a given rational being, using both I and L, a large part of the relevant logical apparatus will be at once part of I and L for the reasoner; use of both I and L will consist in complex attitudes to the relation between this common part and, say, DNE.

Much of this will consist in acceptance and rejection respectively of various beliefs obtained with and without the use of that rule - that is, the reasoner's stock of "truths" and "falsehoods".

6. Körner, (CF), pp. 35-38

3.3 Second Objection : revision is trivial.

An objection to the idea of revision to logic that is almost as strong as the preceding one is this : revision is possible, but its significance is illusory, for it is purely notational. This thesis has been somewhat confusedly attributed to Quine recently, both by Morton⁷ and Haack⁸.

One cannot reject the possibility of revision by simply claiming that everyone uses the same logic "really"; but this at once raises the question : how one determines what logic someone does use? Not to prejudge the issue here is to put the possible deviant in the position of an anthropological object and seek a translation of his logical dialect. The alleged triviality thesis is an argument to the effect that every translation which does not reveal the foreign logic as classical logic must be a mistranslation. Thus

"we impute our orthodox logic to (foreigner)
or impose it upon him, by translating his
deviant dialect."⁹

Thus disagreement is illusory.

"Here, evidently, is the deviant logician's predicament :
when he tries to deny the doctrine he only changes the
subject."¹⁰

For apparent deviants cannot mean the same by the logical connectives :

-
- 7. Morton, A: (DDCS)
 - 8. Haack, S.: (DL)
 - 9. Quine, W. (PL)
 - 10. *ibid* p.82

"surely the notation ceased to be recognizable as negation when they took to regarding some conjunctions of the form $(p \sim p)$ as true, and stopped regarding such sentences as implying all others ..."¹¹

The premise is that the evidence for translating words as logical connectives consists in observing that they obey certain rules of passage, for example

"if a native is prepared to assent to some compound sentence but not to a constituent, this is a reason not to construe the construction as conjunction."¹²

However, as Quine himself points out ¹³, though his critics overlook this, this argument only rules out our identifying a construction with a connective when it contravenes rules governing that connective. Thus adequate translations of foreign theses cannot come out as negations of classical theses. What is not ensured is that

"all our logical sentences carry over into truths of the foreign language ..."¹⁴

In fact, Quine goes on to explicitly consider such cases as intuitionism which fit this description. The notion that he regards these cases as being one of the other two (notational variant or mistranslation) is not borne out by such remarks as

"In repudiating ' $p \vee p$ ' he is indeed giving up classical negation, or perhaps alternation, or both; and he may have his reasons ..."¹⁵

11. *ibid*

12. *ibid* p.82

13. *ibid* p.83

14. *ibid*

15. *ibid*

"It is hard to face up to the rejection of anything so basic."¹⁶

"The price is perhaps not quite prohibitive, but the returns had better be good"¹⁷

Consider the earlier claims I quoted - the first two of this section: is Quine simply inconsistent? I think not. The deviant logician indeed does not deny the doctrine: he has given up classical denial (i.e. negation); and by the same token "the doctrine" must be different in so far as the rules for the connectives in it differ. Yet this clearly doesn't exclude 'conflict' or 'rejection' (cf. the quote above from p.86). As for the very first quotation given, this is I think a mistake - as Quine himself has perhaps recognized in a later work.¹⁸ The points are two: Firstly, we do not necessarily impose our logic on the deviant logician, as Quine unguardedly claims; we may 'impose' only part of it.

If we have decided that a construction is sententially connecting, we determine how to translate it by the methods Quine gives,¹⁹ but these are not of uniform weight. Thus,

"if a native is prepared to assent to some compound sentence but not to a constituent, this is a reason not to construe the construction as conjunction."

might be a weighty reason against that construal if he were actually to dissent from the constituent, for in that case the translation would result in contradiction. But if it is only a case of assent

16. *ibid* p.84
17. *ibid* p.86

18. Quine, (RR), 21
19. Quine, (PL), p.83

withheld, the construction might be translated as conjunction if there is no other candidate and the introduction rule (assent to compound upon assent to constituents) is followed. Then our theory would be that for some reason the elimination rule is not used. It is hard to see why this could be, and harder to see why the construction exists at all in its rather useless form - but not hard to see that this is possible. The case is quite parallel to the example of revisions already suggested.

Prior's²⁰ nightmare of the connective "tonk" which follows one of the rules for "and" ($A \text{ tonk } B \rightarrow B$) and one of those for "or" ($A \rightarrow A \text{ tonk } B$), might seem to refute the foregoing.

The answer to this I think is not to drag in the question of the existence of connectives as does Belnap,²¹ which only adds to the difficult questions before us the yet more difficult ones about mathematical existence. Rather, we can be brutally direct and admit such a connective - why not? - but note that it is useless. It is useless because everything becomes provable - which takes away the point of giving proofs.

(The so called 'problem' of why it's so bad to be able to deduce everything rests on ignoring our need to refute as well as to prove. If proof and refutation lead to the same results, their points are lost.)

20. Prior, (RIT)

21. Belnap, (TPP)

A second point is that $(p \vee \sim p)$ for example is arguably not part of the characterization of any logical connective by its use.²² Certainly one can deny that it characterizes \sim , since for intuitionists it doesn't; and even for Putnam it doesn't enter the "core meaning" of \vee . So if we accept the principle that one identifies connectives by (though not necessarily with) rules they obey, this does not commit us to all of classical logic unless we are implicitly adjoining to that principle the further principle that a foreign construction is not a certain connective of English unless it satisfies all classical theses that connective enters. Naturally, this constraint will yield classical logic, if it yields any!

(Quine seems to have recognized this latter point in the Roots of Reference, where he allows that only "verdict" functions can be identified by his behavioural criteria and that $(p \vee \sim p)$ is a theoretical thesis!!)

All of the foregoing is perhaps superfluous if one can accept that a possible (if undesirable) alternative logic is provided by an implication fragment such as E_3 , which contains but the one non-classical connective.

3.4 Third Objection : Revision Is Irrational

If revision of logic is accepted as non-trivially possible, the next defense against us doing it or accepting that it is done, must be that this would be, inevitably, irrational.

22. Putnam thinks otherwise : cf. (ILE) p.233

We have seen *prima facie* reasons for revision, such as dealing with the epistemology of quantum mechanics, and others are not hard to find - dealing with vagueness, for example. I have also argued that revision would, in some cases, be genuine, not merely notational. Perhaps these reasons could never be good enough ones for changing logic, however. Haack mentions two arguments for this;²³ the argument that revising presupposes logic, and Popper's demand for the strongest logic for criticism.²⁴

Quine indicates a third, still weaker, one in the passages quoted above when he points to the great inconvenience of dropping $(pv \sim p)$. The reluctance of mathematicians to give up the highly aesthetically satisfactory towers and minarets of non-constructive mathematics reflects the same kind of objection. Let us consider these three kinds of arguments in reverse, that is, in order of their strength.

The first kind of argument, the argument from inconvenience, cannot be a knockdown argument, as Quine of course admits in the passages quoted. Rather, the idea is that revision could, logically, be rational; but it is inconceivable how the gains could be sufficient to outweigh the cost. This kind of argument, never very appealing, continually becomes less so as time passes and alternative resolutions for the epistemological quandaries which have led to such suggestions for revision of logic do not become acceptable alternatives to logical revision.

Popper's argument is more substantial. His point is that criticism is strongest if we use "the strongest logic ... at our disposal",

23. Haack, *op. cit.* pp. 35-8

24. Popper, (OK) pp. 305-8

since then the largest class of potential falsifiers for our theories is made available. Putting aside the fact that the strongest logic is an inconsistent one, which shows that other criteria must be involved, it seems that Popper is ignoring the possibility of using classical logic to criticize classical logic, and thus provide an argument for a revision : which is just what intuitionists do.

But in fact, he restricts this argument to the question of which logic we should use in empirical science. Haack argues that his argument here is circular, and I believe that it is; but the circle is not the one she suggests.

According to her it goes

- (i) logical revision would impede science \therefore don't revise.
- (ii) logical revision saves revising some of science (e.g. QM),
 \therefore impedes science.
- (iii) logic is not falsifiable, so not science, so LR saves revising science.
- (iv) (i) \rightarrow logic is not falsifiable.

Now in fact falsifiability is not the same as revisability, though Haack pretends that it is. But anyway, as she admits outside this circle, the reason that logic is not, chez Popper, falsifiable is not because if it were it would impede science, but because it consists of tautologies which conflict with no basic statement : "Every tautology follows from every basic statement, since it follows from any statement whatever."²⁵

Here is where the circularity is, for this claim of Popper's is only true taking "follows from" to invoke classical logic. The formula which Putnam wants to drop is a tautology, as of course is $(p \vee \sim p)$; but if that were enough to make them follow from any statement whatever, they'd be undroppable! Naturally, classical logic cannot be revised while retaining classical logic.

Thus I reject Popper's argument for being circular, but not quite on Haack's ground! So logical revision may be rational after all, as far as this argument goes. If the object against which the "strongest logic" is directed is itself classical logic, we must be prepared to act on the results.

The third argument, that revising logic presupposes logic is only plausible as long as the referents of the two uses of "logic" are kept vague enough for it to be assumed that they are the same. This need not be so. Even if it is, as in the case of using classical logic to criticize itself, the consequent revision, if we decide to make it, does not rest on making use of full classical logic, only enough of that to legitimate a particular inference:

- 1) L has unacceptable consequences (Premise)
- 2) The best resolution to (1) is to abandon L for I (Premise)
- 3) So I should abandon L for I (conclusion)

The premisses may have been obtained using L, but this inference need not : it can be made using I (or even less).

3.5 Fourth Objection : revision can only be peripheral.

The last argument leads naturally to the last and weakest objection to the idea that logic is revisable which I shall consider. This is that although some logical rules may be revisable, not all are. So, we mistakenly thought $(p \vee \sim p)$ could not be revised, but there are nevertheless some other essential parts of logic. One obvious candidate is suggested by the previous argument, namely modus ponens. Similarly, Bennett's insistence on the ineluctability of negation might be seen as arguing the same for modus tollens.

The argument that there is, in a weak sense, a limit to revision, coincides with the argument which excludes more bizarre kinds of deviance which are occasionally hinted at in philosophical discussion of anthropological data, as for example in Winch. Here the idea is that the 'logic' employed by some exotic culture might be so alien to us as to be simply untranslatable. What this means is somewhat obscure - I think in general the claims that nuance X or idiom Y are untranslatable are about as cogent and as important as the doctrine that you can never know what I am thinking ("really") - but one precise interpretation is that the foreign tongue has none of the classical sentential connectives (and, or, not, etc.). This I reject.

Now I do not think it necessary that foreign contain an equivalent for 'and', nor for 'or', nor for any other specific connective. It is possible, though no doubt somewhat unlikely, that foreign should contain only one word for a connective, for example the Sheffer stroke.

But it must contain some connective. My argument for this claim follows.

In the first place, a viable language must allow inferences to be drawn. For if not, every stretch of discourse must be simply a concatenation of unrelated sentences. But in that case what is the point of it? If any linguistic action can be only a response to nonlinguistic action, or a stimulus to it, there seems to be nothing about it to dignify with the name 'logic', deviant or otherwise.

Secondly, if any inferences can be drawn then some sentential inferences can be. (That is, inferences whose form can be given by the propositional calculus.) For suppose not; suppose that in foreign the all and only acceptable inferences are, say, classical syllogisms. So, for example

$$AB^a, BC^a, \text{ so } AC^a$$

is OK. I suggest that a viable language must have a way of separating this discourse from any context of other sentences, to avoid confusion with them. This amounts to a device for conditionalization i.e.

there is some way to put this argument as

$$\text{If } AB^a \text{ and } BC^a \text{ then } AC^a$$

so that it can be discussed. One 'suspends' the premises in order to consider the entailment between them and the conclusion. This does not show that words for 'if .. then' must exist in foreign, but, it does require that some syntactic device so translatable exist. Further, there must be an inverse device for relieving conditionalization ('discharging assumptions') - which amounts to some rule like modus ponens. I think there must be some rule of this kind if there is to be logic at all.

A logic need not contain a negation of any kind - there are for example the 'positive propositional calculus of Hilbert'²⁶, and E_{\rightarrow} the 'pure calculus of entailment' of Anderson and Belnap.²⁷

The question is, whether a viable language could lack all kinds of negation. The basic argument has to be essentially that put by Bennett : reports which relate to the not both present and particular are no use unless they can be challenged. But the reason is not Bennett's reason that this is required to get the concept of rule into the situation. On the contrary, these systems make it clear that the notion of rule is quite capable of employment without such an operator's presence. (ALL logical systems use rules.)

Thus the argument has to be a pragmatic one : could we account for the existence of a linguistic group employing solely the logic E_{\rightarrow} ? Clearly, one way is open to us - we could envisage philosophical arguments leading to its adoption by a group of mathematicians even more epistemologically narrow than the intuitionists, for broadly similar reasons. In the present academic context, such a sub-language could conceivably survive, like intuitionese.²⁸

Since this fantasy is possible, the argument for denial must rely on rationality's involving more than simply concepts and logic. In fact it rests on the demand for empirical concepts, *and so for* semantical rules. Moreover, if these in their turn are to be revisable (cf. Bennett on 'frozen intelligence'), this in turn requires adjudicative rules, as argued in Chapter 2. These might be regarded as logical, or perhaps methodological. In any case, it is

26. Church, (PH) P.140

27. Anderson and Belnap, (E), ch.7

28. There are, indeed, "stricter" intuitionists; cf. Griss, G. (N)

clear that if no specific ones are necessary (as I conclude from Bennett's failure to argue for some), whichever ones are held must be revisable.

Neither of these admissions, however, need commit us to any specific rules to govern constructions we may with some latitude call "negation" or "modus ponens". Once again here we must refer to the general argument set out in chapter one; in fact concerning logic we may insist on it twice : that there are some logical rules is necessary, but of no specific ones can it be said that they are necessary. And more particularly, concerning the notions of entailment and negation, we may say that some rules governing logical constructions we so recognise must exist; but not that of some particular ones, those ones must be accepted by any rational thinker.

CHAPTER FOUR

NAIVE EVOLUTIONARY EPISTEMOLOGY

4.1 Some motivations for evolutionary epistemology

4.1.1 The one thing on which all the more radical philosophers of science agree, and I concur, is that a strictly deductivist account of scientific advance (the "layer-cake model") is wrong. But this admission creates a dilemma. If theories are not deduced from data, if the relation of replaced to replacing theory is not inferability (although deduction still has a central place in scientific explanations), if no a priori account of what is rational in scientific change can be given, then there is no iron necessity in the sequence of theories. Yet to retain some hold on the notion of reason we cannot simply dismiss what is apparently scientific progress as an illusion : in some sense, later theories are (or should be) better than earlier ones. How can we reconcile this direction, with the admitted lack of apodeictic logical directives?

Direction without a director is just the kind of impossibility made possible by Darwinism. The thrust of Darwin's theory is to explain the appearance of design without recourse to a designer. And this requires accepting that the 'appearance of teleology' refers to a real aspect of organisms, later called fitness. In Darwin's theory the fitness (fittedness to its environment) of an organism is explained by a series of improvements, none of which is compelled to occur. Are there any reasons to think that this loose parallel can help with our dilemma? The proof of such puddings is in the eating (cf. Chs. 5, 6 infra.), but there are two reasons for thinking so which, if my thesis is correct, are ultimately one.

4.1.2 The first argument rests in acceptance of the theory of evolution.

In particular, I accept that man evolved from animal ancestors, and that some present day species such as chimpanzees derive from a common ancestor with man. Furthermore, I think that the available evidence about these species, and that on dolphins too, indicates that at least some mental characteristics of man are evolutionary products. Finally, I do not see that there is at present anything but theological prejudice against the view that all human characteristics are the products of evolution, and much economy to be gained from it.

Given this, I take it that the presumption must be against such unexplained lacunae as are required by the notion of a "logical precipice" between man and the animals, or the similar postulation of some "critical point" at which cultural development took over from biological evolution, a punctual emergence of Man. Against the latter I do not propose to argue here, except to ask the rhetorical question of those many philosophers who conceive this watershed as the beginning of "genuine" symbolization : what was the very first message? The absurdity of looking for an answer to this question is evident.

And yet without such a cognitive saltation, how can the notion of a 'logical precipice' be squared with evolutionary theory? I suggest that, as used for example by Bennett, it is a stipulation rather than a discovery, as indeed is apparent from my discussion of Bennett, where I have pointed to the unwarranted (except by this stipulation!) restriction of "reasons" to "conclusive reasons" and "having concepts" to "linguistic capacity". Similarly arbitrary demands are often put on what constitutes language or communication, when these are taken to be the unique possession of man. Where is the evidence for such a difference in kind which does not rest on such a petitio principii?

The phrase "difference in kind" is a rather ambiguous one, carrying a connotation of absolute ontological disparity. But etymologically of course, it refers basically to biological traits, as characterize the difference between, say, pigs and ants. Consider the trait which pigs have and ants don't of being fourlegged. This difference is a real one, but no ontological consequences ensue. The existence of this difference is one of the explananda of evolution theory; the explanation, if given in details, would begin from a common ancestor of all pigs and ants and would provide us with an understanding of just what four-leggedness is through the details of the history of descent to pigs which ants didn't go through, for a cardinal feature of it must be the usefulness to their possessors of various kinds of locomotive system, and reasons why in the circumstances of the proto-pigs, but not those of the proto-ants, a four-legged kind of locomotion was superior.

4.1.3 The second argument for seeking an evolutionary characterization of reason is that a detailed parallel can be drawn between two processes: the process [REDACTED] of biological evolution, and the process of scientific discovery. Most of the writers we are concerned with grant there is an analogy. Popper claims there is more but does not make out his case. I hope to make out a better one.

A summary statement of the parallel would be to say with D.T. Campbell that both processes are examples of the "blind-variation and selective-retention" process.¹ In such a process an initial stock of entities gives rise to a pool of copies of themselves which incorporates a range of (minor) variations, their putative successors; the less appropriate of these are eliminated and the survivors constitute the next initial stock.

1. Campbell (BV) cf. also Wimsatt (TLSF).

There are three basic elements to the evolution theory : heredity, variation and selection. All three are present in scientific thinking : our theories are descended from earlier ones; alternative minor modifications to existing theories are created to deal with problems; not all such suggestions are fruitful (pun intended). But is this identifiable parallel enough for transference of the theory (rather than just the language) of evolution to scientific changes? I shall say something about why it seems insufficient in extant writings in the next section.

4.1.4 There is a third argument for looking to the theory of evolution for guidance for the dilemma of 4.1.1, connected in this case neither with the process in question, nor the subjects/agents of that process (4.1.3 and 4.1.2 respectively). Rather, it concerns the fact that the dilemma arises over giving sense to the notion of 'better than'. Bearing in mind that the dilemma has arisen in the ruins of post-Cartesian epistemology, it might be salutary to consider whether the general notion of evaluation current today does not suffer from something like the same infection as that notion of knowledge.

In particular, I would simply like to point out, without presently developing the point², that on an Aristotelian analysis of (moral) good, the concept of 'better than' has logical connections with the notions of 'well-being' and 'fitness'. Without going to American extremes, must there not be some truth in the pragmatic contention that better theories are of more use?

2. cf. Grene, (UN); Ginnane, (MC) et (many) al.

Having put aside *a priori* presumptions, should not Ockham's razor suggest that we look in the same direction for the source of normativity in general, before we split it between reason and practice?

4.2 Naive Evolutionary Epistemology In The Literature.

4.2.1 There are many signs of an evolutionary consciousness in the literature, but, as yet, no considered and detailed statement of a sound position. The two nearest approaches must be Toulmin's Human Understanding and Popper's Objective Knowledge. The first of these, apart from its offensive vagueness and pomposity has various specific faults, for which I shall largely disregard it; one of them is worth discussing, and will be in the next section. Toulmin's over-riding concern is a misplaced zeal in finding a substitute for the notion of logical systematicity; and the theory of evolution, while definitely concerned with diachronic phenomena which logic (perhaps) isn't, won't provide it. On this aspect of Toulmin, Cohen's review³ is excellent.

One can point to many tantalizing declarations of faith in some kind of evolutionary perspective, and as in Quine (RR), Piaget (BK), Rosenberg (LR), Ackermann (MDL) et al. An excellent bibliography relieves Campbell's article⁴ of its overall tediousness. What most such discussions avoid considering is that the explanatory force of Darwinism derives from its being a theory and the theory must find application, not merely similarities in the phenomena under discussions. An example of this kind of illegitimate "bandwaggoning" is provided by the final few paragraphs of Kuhn's iconoclastic little volume. He declares, in discussing what sense he can give to

3. Cohen, (RT)

4. Campbell, (EE)

the notion of "progress through revolutions", that

"if we learn to substitute evolution-from-what-we-do-know
for evolution-toward-what-we-wish-to-know, a number of
vexing problems may vanish in the process"⁵

a sentiment which I applaud, but of which Kuhn gives no elaboration.

On the next page he is a little more specific:

"The analogy that relates the evolution of organisms
to the evolution of scientific ideas can easily be
pushed too far. But with respect to the issue of this
closing section (progress via revolutions) it is very
nearly perfect. The process described in Section XII
as the resolution of revolutions is the selection by
conflict (sic) within the scientific community of the
fittest way to practice future science..."⁶ (my emphases)

which, while betraying a lamentably Malthusian idea of evolution, at
least fixes on selection as the core of the analogy. But that is as
specific as it gets. And further on it is clear that his appeal to
evolution is purely in the service of the general rationale of section
4.1 above - explaining the appearance of teleology.

4.2.2 Popper's treatment is rather richer of course, although he
is also very vague on some aspects of the metaphor - which, he insists,
it is not.

Despite the subtitle of Popper's Objective Knowledge - An evolutionary
approach - he makes little attempt to make out how the parallel is more

5. Kuhn, (SSR), P.177

6. *ibid.*, p.172

than a metaphor. The evolutionism is in his approach; in particular it bolsters his ontology (cf pp 112 et. seq. 'A biological approach to the third world'). The root of the trouble is his confused notion of the logic of evolution theory. Thus on pp 241/2 he prefaces the most explicit statement of his evolutionistic thesis with the remark that evolutionary theory is "tautological, or almost tautological ... Darwinism and natural selection, though extremely important, explain evolution by 'the survival of the fittest' (a term due to Herbert Spencer). Yet there does not seem to be much difference, if any, between the assertion 'those that survive are the fittest' and the tautology 'those that survive are those that survive' ..."⁷

I shall indicate what is wrong with this view below. (It is common amongst philosophers.) For the present suffice it to say that the fact that Popper holds it may explain his rather arrogant notion that what he gives in the following passage is "a statement (of Darwinism), which makes it less vague".

What he in fact then gives is a phenomenological description which can be applied (or so he claims) both to organisms in their environments and problem-solving in the third world. Even if this is correct it is not enough : Darwinism is a theory not merely a method of description. The strength of present-day Darwinism lies in its positing of genes which enable the mathematics of heredity to provide inferences which help to explain the changes in the description of the biotic phenomena which fossils tells us have occurred (and experiment shows us still do).

In the other main discussion of evolution in (OK), Popper "briefly mention(s) a point or two" on the logical form of a theory of natural selection, regretting that he "should have liked to expound it at

length". It is a pity he did not.

A different aspect of Popper's evolutionary approach is much more to be applauded however. This is the connection he makes between the idea that "each organism can be regarded as a hierarchical system of plastic controls"⁸, with his recognition that the role of rationality in science is that "our discussion is controlled, though plastically, by the regulative ideas of truth and validity." He does not characterise plastic control explicitly, but his examples (automatic pilot, standing upright) and his references to feedback (e.g. on p.241, explaining why our theories do not control us ineluctably (as idealists would have it)), are enough to show that he has in mind the cybernetic notion of homeostasis. I shall argue in the next chapter that the detailed consideration of this notion will provide explicitly a link between the concept of plastic control and the logical form of the theory of evolution. In the last chapter I shall use this link to unpack the consequent evolutionary notion of rationality some more.

4.3 Pressing the Analogy

4.3.1 Although such writings as I have discussed in 4.2 add a good deal of plausibility and force to the analogy of processes I stated baldly in 4.1.3, by adding detail to it, critics of such suggestions tend to rely on the argument that it is, after all, only an analogy, which breaks down if pressed to more detail. Of course, since this pressing tends to rely on a very simple and old-fashioned notion of evolution theory, the conclusion that the progress of science couldn't be evolutionary is hardly warranted, since evolution theory should be, and is being, reformulated to take account of the great strides of 20th

century information-centred science, in particular cybernetics, the genetic code theory and the information-processing approach to the brain. But that is the subject of Chapter five below.

Nevertheless, such criticisms are worth considering, for two reasons. Where confused, they point to aspects of the nature of evolutionary explanation which are commonly misconceived, especially by philosophers. And where not, they serve to indicate ways in which evolutionary theory has not yet been soundly reconstructed along the lines indicated in the previous paragraph. (And if all this makes my general thesis sound somewhat speculative, resting on a theory which does not yet exist, well, it is an essay in philosophy, which can still be distinguished from science - at least, until that theory is established!)

4.3.2 Darwinian evolution concerns populations of organisms, divided into species by the possibility of interbreeding. Can something parallel be said of cognitive entities - theories, hypotheses, concepts? (Or perhaps beliefs, or systems of beliefs? I shall regard theories and hypotheses as abstractions from beliefs and systems of beliefs; I do not intend to stand up for Popper's objective third world.) The abstract requirement is that a relation of relative similarity obtain, as well as some kind of communication gap.

- (i) Concepts. Toulmin wishes to take populations of concepts as the basic analogon of the species - say, the pool of conceptional variants of physics at some time. Jonathan Cohen⁹ has soundly criticized this suggestion on the ground

9. Cohen (RT)

that concepts of, say, physics are not slightly varying copies of one another but essentially different and complementary system-elements - for example 'proton' and 'neutron'.

"This relation between a concept and a rational discipline to which it belongs is, roughly, that of part to whole, not that of member to species or member to population."¹⁰

David Hull has rejoined to this that organisms do not have the platonic relation to each other Cohen implies either.¹¹ But Cohen has retorted¹² that

(1) differences in role (as queens, drones, workers in bees) overlay a deeper identity of genetic structure,

and (2) in "any one version of atomic theory, unequivocally expounded there is only one concept of a proton etc."

- so that Toulmin's analogy makes it a polymorphic species with one of each morph - which is next door to extinction.

I agree with Cohen on this point.

(ii) Hypotheses. Popper frequently refers to the natural selection of hypotheses, for example:

"our knowledge consists, at every moment, of those hypotheses which have shown their fitness by surviving so far in this struggle for existence; a competitive struggle which eliminates those hypotheses which are unfit."¹³

10. (ibid), p.49

11. Hull, (RCT)

12. same issue, p. 334-6

13. OK p.261.

This seems a little more promising, for it is easy to think of sets of slightly variant hypotheses, for example a whole range, OD, of equations suggested for the orbit of Demos around Mars. Indeed this suggests a way of rehabilitating Toulmin's account, if we centre attention on his notion of conceptional innovation, which we might regard as a hypothesis which slightly changes a concept already existing (as in the polyhedron / swan examples above, ch.2).

However, the plain fact of the matter is that hypotheses do not stand or fall alone but in "bodies" - they 'face the tribunal of experience' together, to switch momentarily to Quine's analogy. Clearly, theories are the only real candidates.

- (iii) Theories. Do theories fall into populations? One is tempted to think of theories, platonistically, as each of a kind. Newton's theory is, precisely, what Principia says, no more and no less. Maybe; but there would seem to be a variety of Newtonian theories expounded by different scientists at different times to attack different problems. These no doubt differ in respect of the auxiliary hypotheses they involve, but one could think first of a set which differed only in respect of which hypothesis from the set OD above was included. This case would seem to be fairly close to the wanted parallel. Of course, there are further cases in which what one might regard as basically Newtonian theory has subjoined auxiliary hypotheses of widely different kinds in order to deal with problems of different sorts - say in optics and hydrodynamics. But we might argue that the parallel to bear in mind here would be subspecies in different environments.

Although there is some plausibility in this suggestion, the following point will have to be met : in the example, the 'species' of Newtonian theories, the distinction between the 'hard core' and the auxiliary assumptions would seem not to correspond to anything in the organic case, and yet the former discriminates the species from others. In fact this is not serious, for as Musgrave has pointed out¹⁴, what one would still regard as related theories were floated during the heyday of Newtonianism, which did revise the hard core.

4.3.3 But this question of invariant common components in the theories of a species does raise the question of the next point of comparison, namely, the process of evolution of 'offspring'. In what way is there a "family tree" of Newtonian theories? And secondly, and this is a point often felt to tell heavily against the parallel, are the variations in the 'offspring' random in the same way as is demanded of biological mutations by orthodox present-day Darwinism?

I do not find any difficulty in the notion of a tree of descent of various Newtonian theories. Historians constantly seek to demonstrate sources of one person's opinions in those of another, read or heard. And no-one trying out a new Newtonian theory is going to deny that it is "descended" in this sense from previous ones, whether that scientist's or some others.

14. Musgrave (FC)

Notice, though, that this inheritance is not sexual. This is not a frivolous point, for Mendelism depends essentially on sexual reproduction (contributions of genetic material by two parent generation entities to each of the next ¹⁵), and the mathematics of Mendelism was what saved Darwinism from sinking into disrepute early this century. To transfer any of the explanatory power of Darwinism to the cognitive kingdom will require at least the prospect of a mathematics of transmission which can take Mendelism as a special case. Fortunately, this does not seem unimaginable - the basic point is that the transmission mechanism for theories is (human) language. One may hope that some theory of communication might subsume a version of Mendelism couched in terms of the 'coding' description of DNA-genetics.

What however of the charge, made for example by Cohen¹⁶ that 'mutations' are not random? Toulmin just admits this, and claims that the demand for "decoupling" is not central to Darwinism. Certainly, Darwin himself went some way into Lamarckism, being in ignorance of the mechanism of inheritance; nevertheless as Cohen justly urges, one cannot be too cavalier about this, for the 'Weissmann thesis' (in the form, say, that no information flows from phenotype to genotype - cf. Maynard Smith ch.4) is widely accepted as the "central dogma" of modern theory - though not totally unquestioned. It is in fact not clear what role the Weissmann assumption plays in the basic theory,

15. Fisher, (GTNS), etc.

16. op. cit.

since it depends on a distinction between gene and soma which Mendelism ignores.¹⁷ So the centrality of this requirement is somewhat questionable. (Much of the fervour behind 'Weissmannism' is anti-Lamarckian, anti-vitalist, anti-teleological, rather than part of the positive theory.)

I am not convinced in any case of the other claim of this alleged dilemma, namely, that intellectual innovations are not random. Consider the famous case of Kekule's discovery of the chemical structure of benzene through a dream. It has been argued¹⁸ that, "as someone already working on the problems he already had some reason for picking on the hexagon (the true arrangement of the carbon atoms¹). But this, it seems to me, equivocates on what his problem was. To the extent that he 'had some reason', this must be part of his data, of the problem-situation - which perhaps restricted the set of possible solutions without fixing it uniquely. But then just this element is common to the oversimple description of the problem, and the correlation between the two is a matter of misdescription rather than some mysterious intervention of the Truth into Kekule's ruminations. To make the opposite claim seems to me to deny the possibility of genuinely new knowledge, just as would the absence of random mutations ~~prevent~~ the possibility of genuinely new organic forms.¹⁹

I think this particular issue would be ill-served by a discussion of the notion of randomness, for the literature on that topic is not very enlightening.²⁰ It makes much more sense in Maynard Smith's terms, when it became the platitude that you cannot use the solution to some problem in solving it (an argument Popper has consistently used, e.g. against historicism).

17. cf. the discussion in Waddington, (ed.), (TATB4)

18. Gibson (GK)

19. cf. Campbell (EE) pp. 164-5

20. cf. for example, the discussion in Part III of Schaffner (SCBS)

A further point to notice is that in saying this I am not committed to the idea that selection criteria are quite independent of the sources of intellectual initiative - clearly, the known facts will rule out guesses that ignore them!

4.3.4 There is a more serious charge made, again by Gibson, but by others too, concerning the alleged parallel in the selective parts of the processes. The claim is that whereas the eliminated organic variants are eliminated because they cannot meet their demands of the environment, a new element is involved in the survival of theories : namely, it is required that a new theory be nearer the truth than the old. In other words the survival of theories depends on the reasons there may be for accepting them and this

"introduces a quite distinct kind of causal explanation
not found in biology"²¹

I wish at this point to make only three points, although this question will arise again in Chapter 6, for it is, in another form, the old charge of the 'naturalistic fallacy', the argument that any theory such as mine is self-defeating, destroying any grounds for believing that we have any knowledge, by making its acquisition caused.

The first point is that, contrary to popular belief enshrined in the slogan "survival of the fittest" (which was not Darwin's), the essential point of Darwinism is not survival but differential reproduction : there are still prokaryotic organisms about, one of them might even be the original first organism, yet evolution has nevertheless occurred. Similarly, theories don't have to be

21. Gibson, op. cit.

eliminated (for, Popper indeed, they couldn't be; for me, they can if the ability to understand any record of them is lost); the important point is the process of the "going forth and multiplying", of new theories - i.e. their acceptance by working scientists.

Which leads to the second point : it is not enough that there be good reasons for accepting theory A over theory B. Judgments to this effect must be made. "Being nearer the truth" is no guarantee of survival if no-one thinks so. But this is not in conflict with saying that if it is so, people are likely to think so. The position here seems to me to be parallel to the possible fate of a new strain of, say, marijuanha admirably suited to the conditions in England which are so much wetter than in India - the 'normal' environment, - only to meet an unusually long hot summer and leave no progeny - even though they would have been better in more 'normal' conditions.

(These two points cover the parallel to be drawn to the resources/competition aspect of selection)

My third point is that all this does not bring in a "quite distinct kind of causal explanation not found in biology" because the standard kind of explanation found in evolutionary biology is greatly different from what is generally thought of as causal explanation anyway. But this needs a chapter to itself, namely the next one.

4.3.5 Before discussing these points of detail, I remarked that the difficulties that arise fall into two classes : those due to a misconception of what Darwinism is, and those due to the need to reconstruct evolutionary theory in a systems-theoretic, information-processing form. To conclude this chapter I will summarize the discussion under these heads.

A clear example of illusory difficulty is the population parallel, where implicit Platonism about organisms, contrary to Darwinism, may involve an apparent problem; and a clear example of a genuine difficulty raised by the same point of comparison is the obscurity of the role of the gene/soma distinction in evolutionary theory, which makes its cognitive analogon inevitably obscure. The objection resting on the Weissmann dogma, the randomness of innovation thesis, is partly of one kind and partly of the other : in so far as it is not clear how essential to the theory, it is a matter of reformulation of the theory but in so far as it is not an essential assumption (merely a mathematically helpful assumption for Fisher's calculus, say), it embodies a misconception of evolution theory. (That it may rest on a mistake about the data about intellectual innovation is a third possibility, as I argued.)

Two more points which arise from misconception concern the notion that selection concerns survival (which is just false), and the notion that an improved variant must survive (the theory is statistical at heart); and another point indicating a need for a reformulated theory is my own objection that cognitive descent is not sexual - this points to the need for a more generalised theory of the transmission of stored information/instructions.

Finally, the objection that evolutionary explanation is causal whereas rational explanation is something else, is firmly under both headings. This is the most serious difficulty and in trying to deal with it in the next chapter, the others will, hopefully, fall into place.

5.1.1 Introductory remarks

A common and apparently potent objection to an evolutionary account of scientific change is that it neglects the role of rationality in the process. It is claimed that evolutionary accounts of scientific change are "irrationalist" (Lamarckian) in that they ignore the role of rationality in the process. This is a serious charge, for it is one thing to say that evolutionary accounts of scientific change are "irrationalist" in that they ignore the role of rationality in the process, and quite another to say that they are "irrationalist" in that they ignore the role of rationality in the process. The latter claim is a much more serious charge, for it is one thing to say that evolutionary accounts of scientific change are "irrationalist" in that they ignore the role of rationality in the process, and quite another to say that they are "irrationalist" in that they ignore the role of rationality in the process.

CHAPTER V

THE LOGICAL FORM OF EVOLUTIONARY EXPLANATION

5.1

5.1.1 Introductory remarks

A common and apparently potent objection to an evolutionary description of scientific change is that explanations of rational behaviour or belief ("rational explanations") are different in kind from explanations on non-rational phenomena ("causal explanations"). This need not be taken to include the claim that there is only one kind of causal explanation - explanations of animal behaviour obviously differ (at present anyway!) from explanations of the motion of billiard balls. But there is alleged to be a gulf dividing rational explanations from others.

In the previous chapter I suggested that, as an evolutionist, one must demand reasons for positing such gulfs. Reasons offered might concern the content or the form of rational explanations. The simplest such move, positing intelligibles, has recently been argued again by Popper, but apart from the traditional objections to Platonism, suffers from begging the question : have non-humans access to the third world? If so this move will not provide a gulf; if not, we are still in need of an explanation of how proto-humans could ever have crossed the gulf.

But perhaps the characteristic constituents of rational explanations are not propositions so much as constructions like "X's grasping that p" which have alethic, hence normative, implications ...

But can we not with propriety speak of a dog's grasping that he is in danger? To call this anthropomorphism is not only in conflict with evolutionary explanations of, say, flight behaviour, but again a petitio principii.

Finally what of the normative implications : that one can say that X's decision was rational because the evidence was thus-and-so and so he should have believed p. But cannot one make precisely parallel remarks about our threatened dog? Is there not here too a distinction between cases of justified believings-in-danger, and others not really backed up by the evidence (master feigning anger, film of elephant stampede, etc.)?

These brief remarks must suffice to clear the way for a consideration of the alternative argument, that in some way rational explanations differ in form from causal ones, and so couldn't be evolutionary explanations. My strategy is to argue in this chapter that evolutionary explanations are not causal in the required sense either. In the next chapter I try to apply to reason the characterization of evolutionary explanation which I propose. On the way we shall come across further reason for rejecting the notion of ingredients of odd status in rational explanations.

5.1.2 Context : Extant Literature

Discussion in the recent literature as to the nature of evolutionary explanations is neither extensive nor helpful. It all revolves around the question whether such explanations conform to the covering law model of the Hempel/Nagel school. Thus Goudge¹ argued strongly for the necessity for narrative explanations, which he connected via the image of crossword puzzles, to the idea of rational explanations as found in history, and argued by Gallie, Dray and others to be not Hempelian. Similarly, Scriven² argued that evolution theory fails to measure up to the Hempelian demand of explanation/prediction symmetry.

1. Goudge, (AL)

2. Scriven, (EPET)

Against them, Lehman³ and Ruse⁴ are concerned to argue that, on the contrary, evolution theory does conform to the covering-law model. (The inconclusiveness of this dispute will be explained by my own account, below.)

I do not propose to enter this controversy, not really believing in the Hempelian doctrine. In any case, I am prepared to grant the covering-law claim, at least in the weak version that evolutionary explanations involve inferences among whose premises are generalisations and singular statements ("laws" and "initial conditions"). The distinction I wish to draw is within the class of explanations so conceived. Inferences of the same form at one level of analysis may differ at another - as do Barbara and Celarent, both of the form $\langle p, q, r \rangle$.

Rather more light is cast on the nature of evolutionary explanations, in an indirect way, by recent discussions of functions and functional explanations. The long-standing puzzle here is to give an account of the explanatory role of statements like

(HF) The function of the heart is to circulate the blood
which cannot be replaced by the metaphysically harmless

(HC) Hearts circulate blood

without vitiating its explanatory powers, even if HC is expanded to a "hypothesis of self-regulation" (Hempel⁵), that is, a cybernetic account of how hearts work.

That kind of systemic⁶ account of functional statements is

3. Lehman (FEET)

4. Ruse, (PB)

5. Hempel (LFA)

6. Cummins, (FA)

confronted in the literature by the etiological⁷ view which tries to revive the explanatory force of (HF) by adding to (HC) the claim (OH) Hearts exist because (HC).

The major difficulty with this view is to give a sense to "because" which makes (OH) true : It is hard to avoid explicating it in such a way that it leads to the attempt to "deduce hearts from circulation"⁸, which you cannot.

After my discussion of the nature of evolutionary explanations, which was arrived at by considering the defects in the two views of functional explanations just outlined, I will show how it can dissolve the problem they pose.

5.2 The evolutionary explanation

5.2.1 The variety and complexity of phenomena which evolution theory attempts to explain is enormous : firstly the 'story' of evolution - that all life has a common ancestor, the increase in numbers of organisms and species, the increase in size and complexity of organisms, the long-term improvement in their efficiency, the spreading of life into the different environments that exist (sea, land, desert, snowcountry; trees and caves); the extinction of most species; phyletic evolutions such as that of the horse; the original problem of speciation; certain key episodes such as the origin of multicellular organisms; taking to the land; the derivation of birds from reptiles; and so on. On top of all that are the phenomena we can presently observe around us, all the complex detail of the diversity of life, the amazing adaptedness of organisms to their environment, their geographical distribution and

7. Wright (F), Ruse (PB) etc.

8. Cummins' phrase.

the system of their structural features, their similarities and variations - the homology of arms, wings and flippers, the uselessness of the appendix, and so on and on and on.⁹

5.2.2 Actually, the foregoing is a trifle inaccurate, as well as unsophisticated, as a description of evolution's explanandum.

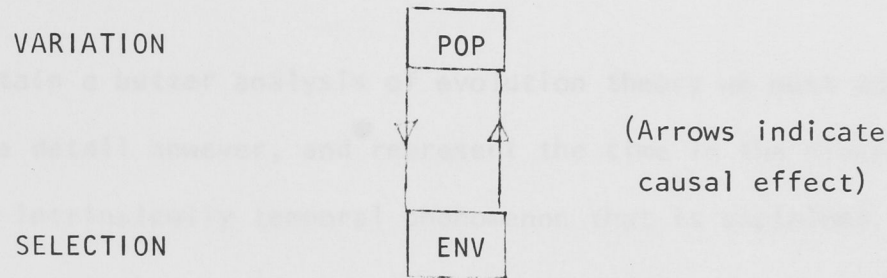
There are two points I wish to draw attention to at this point. In the first place, it is important to realize that much of the so-called "story" of evolution is a theoretical reconstruction, for apart from a few unreliable accounts of observations from antiquity, the phenomena to be explained consist simply of the characteristics of presently living creatures. This is slightly inaccurate in that fossil evidence is important, and that the corpus of "natural history" fieldwork is too; but it is fair I think to say that the fundamental explanandum is the enormously complex pattern of distribution of "adaptive" characters amongst the set of organisms on earth at a given moment.

In the second place, regarding the sophistication of the description : I think it is important to recognise that my phrase in the previous paragraph "complex pattern of distribution" should be taken seriously. It is a gross misrepresentation of evolution theory to suggest that it attempts to explain why there are tigers, and why there are mosquitoes, and; whereas the real aim is to explain why there are (tigers and mosquitoes and ...). The last bracket is intended to point to the unity of the explanandum, which is obvious enough from other points of view : for example, the assumption of a single common environment; and the article of faith amongst most theorists of the

9. This list is partly derived from Goudge.

essential unity, from its common unique origin, of all life.¹⁰

5.2.3 The basic, Darwinian, theory of evolution explains all this in terms of a causal cycle between two black boxes, thus:



POP is a vector consisting of the genetic distribution of organisms of a species at a particular time; ENV is another vector characterizing the environment. Components of ENV would indicate numbers of organisms in ecologically related species (predators and prey, etc.) as well as variables describing the physical environment. For Darwin, the box POP was almost black; by a process quite unknown to him, a population changed its own numbers and genetic constitution (by reproduction and variation); nowadays we have Mendelism and systems-analysis to make it much less so. But of course he did assume increase, à la Malthus. As for ENV, the processes here are known to be statistical and to involve such life-events as being eaten, drowned or buried in a landslide, finding food or a mate and so on. In such broad terms this box was grey rather than black to Darwin. Notice that the basic model consists of a causal cycle - changes in POP, via ENV, affect POP.

The paradigms for a causal explanation I take to be of the transference of motion between billiard balls and of falling bodies; other paradigms of causal connections are the famous 'strike a match and it lights (sometimes)¹¹ and windows shattering when struck. A case which

10. cf. for example, Dobzhansky (GEP), ch1.

11. cf. the discussions in Sosa, (CC)

may not be causal in this sense in Hempel's nonstarting car (it depends on the wiring diagram!). I repeat that the distinction is relative to the analysis as logical form always is. A chain of causes counts as causal for my purposes, but loops don't!

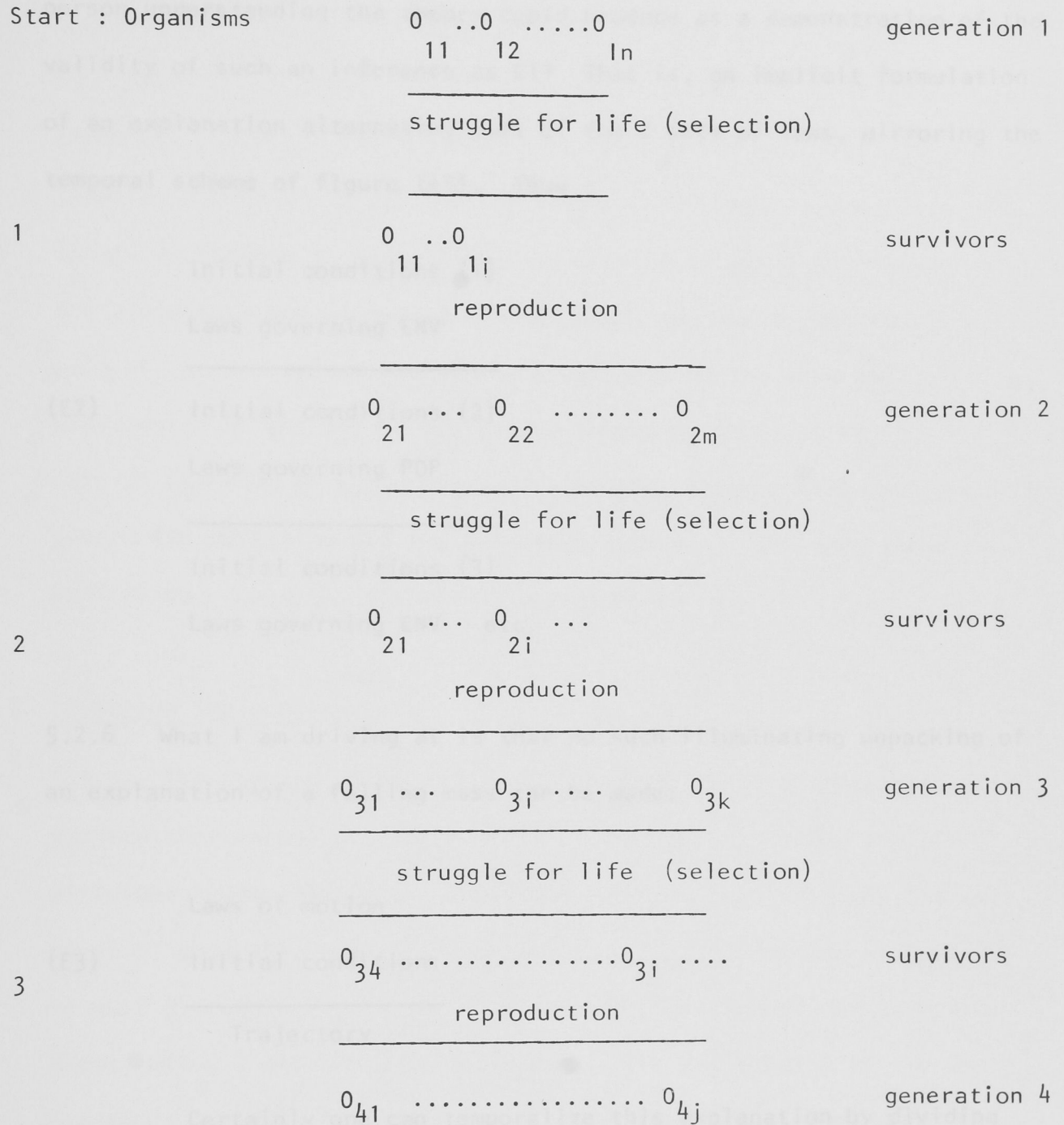
5.2.4 To obtain a better analysis of evolution theory we must add a little more detail however, and represent the time in the diagram : for it is an intrinsically temporal phenomenon that is explained.

So the following diagram is perhaps slightly more illuminating.

Let us regiment the explanandum lightly in the interests of clarity : assume an original population of organisms leading very regular lives - after one time unit, they reproduce (in varying fecundities) and promptly die (altogether); one time unit later what remain of this second generation do the same (there are forces of attrition which prevent some getting to the reproduction stage); a third time unit later, the survivors of their progeny, and so on. cf. the diagram on the next page.

5.2.5 The virtue of this scheme is that it emphasizes the temporality of the explanandum of evolution : it is essential to it that the two kinds of phase of it, the reproduction and selection phases alternate. This point is obscured by a standard covering-law formulation of the explanation:

	Laws governing POP
(EI)	Laws governing ENV
	Initial conditions
	<hr/>
	history



and so on

Is this the explanation, or is it rather one that an intelligent person understanding the theory could produce as a demonstration of the validity of such an inference as E1? That is, an implicit formulation of an explanation alternating uses of the 2 sets of laws, mirroring the temporal scheme of figure (ii). Thus :

	Initial conditions (1)
	Laws governing ENV

(E2)	Initial conditions (2)
	Laws governing POP

	Initial conditions (3)
	Laws governing ENV etc.

5.2.6 What I am driving at is that no such illuminating unpacking of an explanation of a falling mass can be made:

	Laws of motion
(E3)	Initial conditions

	Trajectory

Certainly one can temporalize this explanation by dividing the time, thus

	Laws
(E4)	Initial conditions (1)

	Trajectory (1)

Laws

	Initial conditions (2)

	Trajectory (2)

(Drawn from Trajectory (1))

where the sub-trajectories fit together sequentially to form the whole explanandum. Yet this case clearly differs profoundly from the previous one. Where they differ is in the fact that the order of the subexplanations in the case of evolution is part of the original explanation, but in the physical case this is not so; for that, the temporal dimension is all that is required for their proper sequencing. Second difference : the Galilean explanation is uniformly decomposable - the trajectory may be cut up at any point. And most importantly, each of the subexplanations E4 is an example of the same kind of explanation as E3; but no one of the subexplanations of E2 is a case of the kind of explanation E1. Simply, because different processes govern the variation and the selection phases. The next subsection will discuss what the difference consists in.

5.3 Selection Explanation

5.3.1 The argument just made depends entirely on the thesis that the explanations of the reproduction and selection phases are of different kinds. In fact I shall argue that although we may, at this level of analysis, regard the explanation of reproduction as causal, we must distinguish from it the kind of explanation of the selection phase, which I wish to call cybernetic. The explanation of the whole reproduction phase is simply the aggregation of the individual explanations of reproductions by individual organisms (we ignore sex for simplicity). But there simply do not exist such individual explanations of why an organism does or does not survive the whole selection phase. Rather, a statistical explanation of the set of survivors is provided by a probability distribution for the class of all possible sets of survivors.

The causal explanations in question here are those which apply to the reproduction phases of evolution. Until recently, this was not very well explained at all, except in the gross terms of the "facts of life", but that is sufficient at this point of the argument. It is apparent that the explanation of how a new organism comes into being can be given in terms of sexual activity, pregnancy and birth (for mammals), and that there is no causal cycle involved - there is no reason (in general) why parent organisms need survive parturition, and so need not even exist to be causally acted on by the offspring. Variations are of course assumed to be random, which in this connection implies unexplained, while differences in fertility are assumed to be simple causal consequences of (unknown) genes.

(I am aware that the modern explanation of the mechanism of reproduction does introduce causal cycles in its description of the genome as a system under hierarchic control.¹² But on the one hand, the heart of evolution theory is Darwinism which has no explanation of inheritance; and on the other, the account of evolutionary explanation I am developing may very naturally be extended to integrate 'the DNA story' without significant change.)

5.3.2 The statistical 'explanation' is explained by the use of functional statements. (Statistical explanations are rather doubtfully explanations at all; I prefer to think of them as descriptions of a population ignoring certain details, with various mathematic consequences.)¹³ As a crude illustration of the connection we might contrast subpopulations of some protoamphibians with and without rudimentary lungs. The relatively greater success

12. cf. e.g. Monod (LL), Waddington (SG), etc.

13. On statistical explanation, Salmon (SRSE) is good; for redescription as a necessity in biological explanation, cf. Burhenn (NER)

of the former group at the sea's edge is explained by "the function of lungs is to obtain oxygen", "the function of gills is to obtain oxygen" and the physical facts about the way oxygen occurs in different forms in the sea and out of it. In one kind of environment, the one kind of structure achieves the function better than the other, and the corresponding strain of organism flourishes. (This kind of consideration gives the lie to the claim that evolution is tautologous : lung-equipped animals are "fitter" out of the sea not because they survive but because lungs are fitted to obtain oxygen from the air while gills are not).

In practice it is extremely difficult to put a figure on these things of course, but in principle one may regard the conditional survival probability of organisms with a certain character as determined by the proportion of environments in which that character can fulfil its function out of the total possible environments. Of course, the probabilities for all the different characters of an organism combine in a very complex way, and nothing can usefully be said about an individual's survival probability. Nevertheless a consideration of the way some character fulfills its function will show how, in principle, these probabilities are determined.

5.3.3 Now we must realize that an organism is an adaptive control hierarchy. That is, it consists of a very complex, hierarchically ordered system of regulatory mechanisms which "functions properly" as long as certain crucial variables remain within certain tolerances - obvious examples are body temperature and oxygen supply to the brain. Of course these are only gross descriptions of more complex requirements, and there are many others, some subtler, some grosser (like neck diameter!)

(I think of strangulation). In order for these values to be maintained a very complex network of homeostatic mechanisms is employed - as for example sweating and shivering compensate for changes in the ambient temperature. It might plausibly be argued that in any organism in which gene and soma can be distinguished, the function of somatic characters must be to insulate the essential genetic process of reproduction from adverse changes in the environment and so form part of some homeostatic system. But it suffices for our purposes to observe that this is so for the vast majority of phenotypic characters at least. (Very recent evolutionary work¹⁴ allows that some characters may be neutral to selection; but not all could be). And so long as some are, my claims about the nature of the explanation of the selection phase will stand.

I am not trying to claim that a character like wing colour in moths is 'really' part of a homeostatic system. Biologists, indeed, distinguish such traits from those of which I have been speaking.¹⁵ But one can argue that having such adaptive characters as colour entails having homeostatic ones too. One argument is that traits such as wing colour are clearly in a sense accidental to the wing which exists not to be some colour but for more directly life-support reasons (locomotion, to get to food, to maintain the digestive variables!). Why would an organism have a body consisting only of such accidentally adaptive traits as wing-colour? Is it even conceptually possible? A second argument is that a non-homeostatic character with a certain function is at a selective disadvantage to a homeostatic one - because the latter is more specific to the function. Consider thick skin as an alternative to sweating/shivering

14 cf. Lewontin (GB)

15 cf. Maynard Smith, (TE), ch.1.

in maintaining internal temperature. Thick skin cuts the organism off from information that might be useful (for hunting, say) in order to achieve one specific purpose; but a regulatory system does not.

5.3.4 What is the nature of these homeostatic explanations? I shall discuss a simple example, and argue that the explanation is not causal. This will be sufficient to show that selection explanation, and so evolutionary explanation is not causal. (I am not asserting that all cybernisms are homeostats, by the way, nor even that the cybernetic element of evolutionary explanation is; indeed, it seems clear that many positive feedback mechanisms are involved too (the notion of 'growth' finds its explication here) and arguably the process as a whole requires such a characterisation). Since an organism has been selected out if some vital function breaks down, the explanation of the selection phase must involve this functional explanation.

A standard example of cybernetic explanation is provided by a room with a heater controlled by a thermostat (cf. fig. (i)). This brings out the essentials of the sweating/shivering case without physiological irrelevancies. The room is not perfectly insulated, so heat may pass in or out depending on the relative temperatures inside and out. The internal temperature is measured at T , and depending how it compares with the goal temperature (say 20°C), the switch S turns on or off the heater, H . The purpose of the system is to keep the temperature "close to" the goal value. If it is working properly, a graph of the temperature over a period P might look like figure (ii).

What might one want to explain here? Well, one obvious thing

is the fact that

- (1) During the period P, the temperature remained close to 20°C despite fluctuations in the weather

Notice that this is not the same as

- (2) During the period P, the temperature followed the graph of figure (ii).

Nevertheless, we can explain (1) by explaining (2), as long as we accept the enthymematic premise that following that graph is a case of "remaining close to 20°C ".

Figure (i)

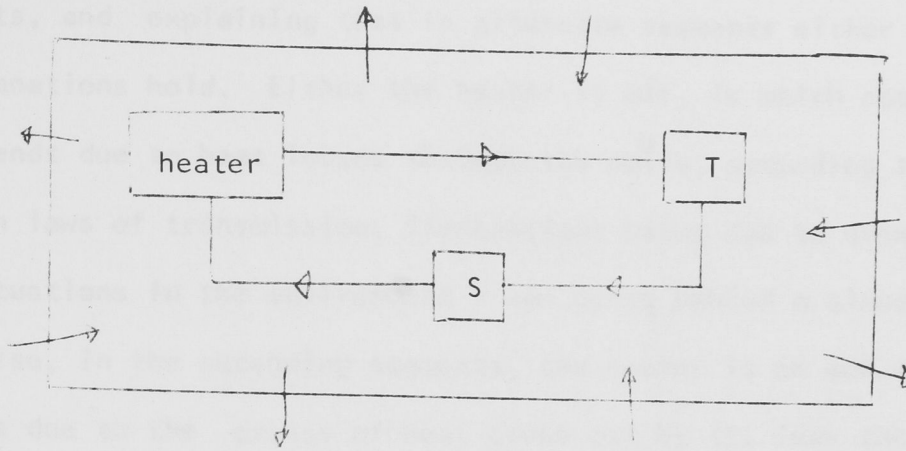
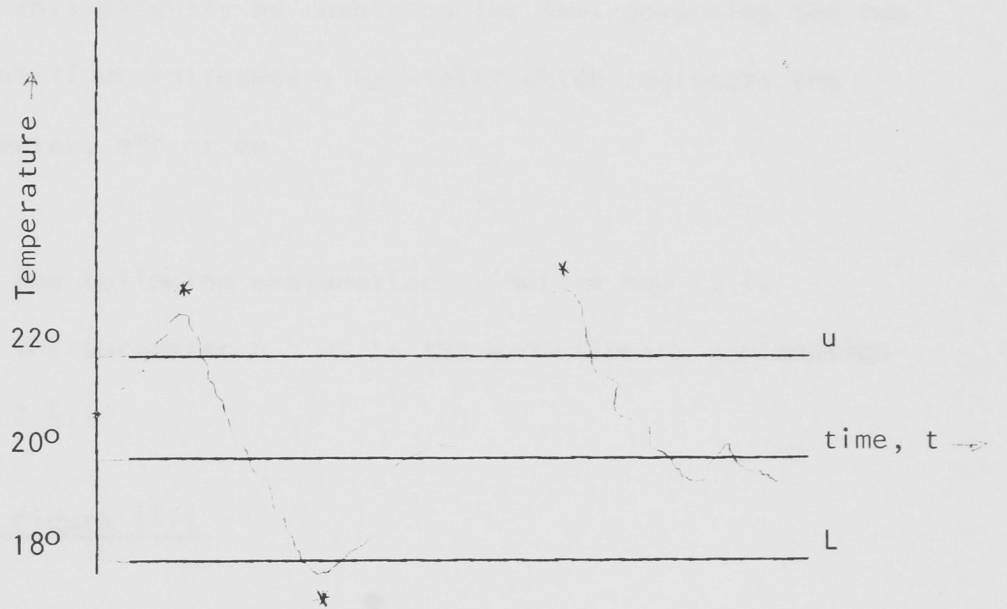


Figure (ii)



An explanation of (2) is given by dividing it up at the starred points, and explaining that in alternate segments either of two explanations hold. Either the heater is off, in which case the graph descends due to heat losses through the walls, according to well-known laws of transmission, fluctuations being due to unexplained fluctuations in the environment - sun going behind a cloud, etc. Or else, in the ascending segments, the heater is on and the temperature rises due to the excess of heat given out by it, less the losses through the walls.

We can improve this slightly by combining the laws governing the two kinds of segment if we introduce a parameter which registers the state of the heater, off or on.

Thus we obtain the following explanation. (Notice how it is articulated by the parameter S - as is the evolutionary explanation discussed in 5.2.4.).

Explanation of figure (ii)

- (1) Over the period in question, the temperature is given by a certain function

$$T(t+1) = f(T(t), TA(t), S(t))$$

where $T(t)$ is the internal temperature at time t

$TA(t)$ is the external temperature at time t

$S(t)$ is a parameter which is 1 when the switch is on

0 when the switch is off

Adding to this the initial conditions

$T(0) = t$ (initial temperature of room)

$TA(\underline{t}) = \underline{T}$ (a certain vector of temperatures)

$S(\underline{t}) = \underline{S}$ (a certain vector of switch values)

The way is now clear for the deduction

(#1)

L $(T(t+1) = f(T(t), TA(t), S(t)))$

Co $(T_0 = t)$

C1 $(\underline{T} = TA(\underline{t}))$

C2 $(\underline{S} = S(\underline{t}))$

 $T(\underline{t}) = \underline{B}$

where \underline{B} is the vector of temperatures as in graph, fig. (ii).

(This is definitely a deductive-nomological explanation)

(2) However, this explanation can be improved if we add an explanation of $S(t)$:

thus $S(t+1) = S(t) \left(\left[\frac{U}{T(t)} \right] \right) + \left(\left[\frac{L}{T(t)} \right] \right) (1 - S(t))$

So the condition $S(\underline{t}) = \underline{S}$ may be replaced by this further law. U and L are the tolerance limits of $T(t)$ - cf. figure (ii).

The explanation now is

(#2) L1 $(T(t+1) = f(T(t), TA(t), S(t)))$

L2 $(S(t+1) = g(T(t), S(t)))$

Co $(T_0 = T)$

C1 $(T = TA(\underline{t}))$

$T(\underline{t}) = \underline{B}$

Now let us return to the task of explaining (1) - remember, we only want to explain (2), that is, the specific graph of figure (ii), as a means to explaining what puzzles us about this room, which for now

we are taking to be (1). No particular interest attaches to why that particular graph should have been followed, rather than some other graph which would also count as a case of (1).

(1) is rather vague and clearly needs to be spelled out, however. There are two points. One is that some kind of bounds need to be specified for what properties of a graph constitute "remaining close to 20°C". This will take the form of a probabilistic specification, such as that there should be no more than a 1% chance of the temperature falling outside the range 18° - 22°. Various further complications are possible here, such as that the average temperature should be 20°, to within some tolerance, but let us leave them aside for the moment. We can now supplement the explanation (#2) with the claim

$$K(T(t)) < 0.01$$

which conveys the fact that (2) is a case of (1), and we now have an explanation of (1). (K is a measure of the spread of T(t), determined by the probability specification).

However, I do not think that we have yet arrived at an example of a cybernetic explanation. The reason for this is quite simple. A statement of an explanation of (1), which is adequate to how we can see the system works, must support counterfactuals. That is, it must show not only how this graph came about, but also satisfy us that had the day been cloudier, or the external environment varied in some other way, then some other graph, also such that (1) was true, would have been produced instead. What we so far cannot explain is the way in which environmental vagaries are systematically combatted by the operation of the system. As Ashby has it, "Only variety can destroy variety"; and we need to make plain how there is sufficient variety available in the system

to maintain (1) in the face of any environmental conditions (within limits of course - every system can be overtaxed). In short, we are still lacking an explanation of the capacity of the system for behaviour of the kind mentioned in (1). Only then will we have an explanation of the statement

(3) The function of the thermostat is to maintain the temperature close to 20°C.

How can we provide this? In fact the method is already implicit in the shift from explanation #1 to explanation #2; what we need to do is to give the canonical representation of this room as a state determined system.¹⁶ To do so we must provide a transition function for the variable $TA(t)$ as we did for $S(t)$. Now of course we can't do this, in a sense, since we are taking the environment as given. Still, it is an assumption implicit in what we have been saying, which we must now explicate, that the range of environmental vectors which would also produce a graph of the kind we want contains the range of environmental vectors we are at all likely to get. We specify this by stating a probability distribution for the difference

$$(SC) \quad TA(t+1) - TA(t) = \text{p.d.}$$

with specified parameters and form, which will be based on known weather patterns, etc. (Assume for the sake of argument that it is a normal distribution with standard deviation σ)

Now the fact at which I am trying to get is this : this probability distribution states among other things, that increments larger than a certain value, say 1, are very unlikely. More important,

16. cf. Ashby, (DFB)

it must imply that the heat transport through the walls is dominated by the heater's emission. This will be expressed by a certain relation between the various parameters of the system, including σ . That is, it sums up when the system will function and when it won't.

Of course, if we make this replacement, we no longer have the specific TA vector and the specific graph of figure (ii) is no longer explained. On the other hand, the explanation of (I) is much more satisfactory this way, being more general. *In fact it is a classic case of subsuming the particular under the general!*

Note that the explanation is now definitely statistical.

It is not often made explicit (e.g. not by Nagel ¹⁷) that this is unavoidable.

How does the new explanation, explanation #3, which is #2 with SC replacing C1, support counterfactuals? Quite simply, by (partially) limiting the set of possible worlds. It does this by the statistical clause (SC), which implies that only in very exceptional circumstances would *the environment be so extreme that there would result graphs not counting as (1)*. So the kind of counterfactual is this :

(CF) S because V; moreover, if not V but $V \cong V$, then S.

whereas a simple causal explanation "E because C" will only support the weaker counterfactual :

(CF') if not C then if C then E.

(The relation \cong is described by the statistical environment limitation).

17. Nagel, (SS), ch. 12

What features of this explanation distinguish it from a causal one? Whatever they are must flow from the closed loop and nothing more - all the rest of the explanation (assumed, not actually given here) is causal enough. I suggest the following features are important enough to jointly constitute a distinction of kind.

Firstly, the explanation supports a kind of counterfactual which a simple causal explanation will not. This indicates that there is a difference of logical form.

Secondly, what is explained is a capacity of the system rather than an event. This underlies the first difference and is itself explained by the statistical characterisation of the system and its environment, which shows that functioning of the system will allow a mapping of environments to system states (as statistically redescrbed).

Thirdly, the explanation is formal : how thermostats and furnaces work is quite irrelevant. The explanandum is the constraint on system behaviour (of not getting away from 20°) and the basic explanans is the form of the causal network (negative feedback). It is in fact an explanation of information control, in the strict sense of Ashby¹⁸. For what is explained is the fact that the environment cannot entirely inform the system with its temperature fluctuations : randomness is being destroyed. That is, information is being controlled.

18. Ashby, (IC)

5.4 Functional Explanation and Evolutionary Explanation

5.4.1 So, cybernetic explanation is not causal, for, among other reasons, the explanandum is the dimensionless quantity information : what is explained is why the survival probability distribution over the set of sets of organisms with such features differs from a purely random one. What explains it is (very largely) the cybernetic account of the functioning of organisms. Where then does this leave the vexed question of functional explanation?

I suggest that it leaves it nowhere : there is no such thing. Consider the alleged functional explanation of my heart. Obviously, what my heart does cannot explain its origin (future causes are absurd), although it can and does explain (partly) its continuance - I need it to keep going, which is a prerequisite for its keeping going. No, the explanation of its origin is this : the roles played in their respective systems ("functions" - cybernetic sense) by the hearts of all my ancestors explains why my particular one exists. For the explanation of there being a genetic network leading to me must appeal to the "function" of all those pre-Coleman hearts which contributed to the survival to reproduce of my ancestors. What causes muddle is using "the heart" to refer indiscriminately to both mine and theirs.

But this explanation, of course, is merely a part of the evolutionary explanation (or an example of one, if you will). Actually it goes beyond explaining my heart's existence in terms of its function - as it must : for the theory of special creation is, prima facie, a possible alternative. A "functional" explanation must rely on a theory of origins, and those we regard as any good rely on the theory

of evolution. My claim is that evolutionary explanation consists of causal and cybernetic explanations, suitably articulated; and that there is neither need nor room for any additional "functional" explanations.

5.4.2 It is a corollary of the previous subsection of course that the explanation of selection phases cannot be causal, being composed of a large number of cybernetic explanations (and some others perhaps). Moreover, the composition of them is clearly not a simple aggregation, for the various adaptive cybernisms of organisms themselves form a hierarchy of systems, and the organisms are of course in a further ecological hierarchical system.

Consequently, further, evolutionary explanation, articulating as it does a complex alternation of selectional and reproductive explanations, (remember that 5.2.2 is a very regimented description!), is not causal. This is the major conclusion I wish to draw from this chapter. But I will go on to try to state what, positively, is the nature of evolutionary explanation in anticipation of my final thesis, that the rationality of science has an evolutionary explanation.

Evolutionary explanation is an explanation by form of an information control hierarchy. If we look at history from the egg's point of view¹⁹, we can see that the essential task of evolution theory is to explain the noisy transmission over time of a message : the message being the information structure of the characters of living things - the original message having been the structure of a (probably) unique first self-replicating macromolecule.

19. 'hens are a way to reproduce eggs' (S. Butler)

Obviously the DNA explanation of the mechanism of heredity can only help my account : it naturally leaves the two part structure of evolutionary explanation untouched, in that causal explanations are still involved at lowest, biochemical level. But yet more layers of information storage, transmission and control are introduced into an already inevitably complex structure.

CHAPTER SIX

EVOLUTIONARY RATIONALITY

My aim in this final chapter is to sketch an evolutionary account of rationality; my main contention will be that rational explanation is an example of evolutionary explanation. This would ideally be buttressed by a history of evolutionary biology, but as it turns out I find evolutionary evidence on rationality unconvincing (an empiricist's dream of thought-experiment?).

By way of preparation, I must first justify the title.

In my first sentence, namely that to show that rational explanations are evolutionary explanations does provide a satisfactory

satisfactory "account" of rationality. The first point is that I

identify a rational explanation with the explanation of a rational

phenomenon. Rationality is a property of phenomena, not of

(or even complementary) explanations. The same thing

Sometimes we loosely describe someone's explanation of his actions

as rational with an implicit contrast to possible (irrational)

explanations. For example, Nikolai Levin in Anna Karenina accepts

extreme action to make Kitty feel better, rather than to save his

soul. But here I think we are misled: to describe this as

mean that the right explanation shows the phenomenon to be rational,

rather than irrational. Whereas the wrong explanation can give the

illusion that the phenomenon is rational when it isn't.

Nevertheless, one might wonder whether explanations of rational

phenomena add up to an account of rationality - do they tell you what

it is? Would you be able to tell me what it is better to

better? I think not, for the following reason.

6.1 My aim in this final chapter is to sketch an evolutionary account of rationality; my main contention will be that rational explanation is an example of evolutionary explanation. This would ideally be buttressed by a history of rationality based on evolutionary data; an enormous undertaking, but I am sure it can be done (an empirically sound substitute for Bennett's thought experiment).

By way of preparation, I must first justify the claim implicit in my first sentence, namely that to show that rational explanations are evolutionary explanations does provide a philosophically satisfactory "account" of rationality. The first point is that I identify a rational explanation with the explanation of a rational phenomenon. Rational explanation is not one among several competing (or even complementary) explanations of one and the same thing. Sometimes we loosely describe someone's explanation of his actions as rational with an implicit contrast to possible irrational explanations : for example, Nikolai Levin in Anna Karenina accepts extreme unction to make Kitty feel better, rather than to save his soul. But here I think we are misled : in cases like this we mean that the right explanation shows the phenomenon to be rational, rather than irrational. Whereas the wrong explanation can give the illusion that the phenomenon is rational when it isn't.

Nevertheless, one might question whether explanations of rational phenomena add up to an account of rationality - do they tell you what it is? Would not an analysis of the concept (à la Bennett) be better? I think not, for the following reasons.

Firstly, 'what-is-X?' questions have been notoriously resistant of answers except by the stratagem of theoretical explanation as found in science, which one might say is by contrast with 2000 years of metaphysics ('what is being?') wildly successful. Secondly, humility and experience both point to the impossibility of a satisfactory a priori, armchair analysis. We simply do not know what rationality is, not because the knowledge is locked away inarticulate in the depths of our linguistic practices ('what would you say about hive three, 0 native speaker?'), but because we just don't know. The only way we can find out anything is by guessing under critical control : theorizing and experimenting.

Thirdly, even if an analysis were achieved, would it be a satisfactory account anyway? (now we have explanations of other things as benchmarks for accounts). What could an analysis do but provide a "reminder" that reason is not a substance, which is a philosopher's red herring anyway, while leaving its possibility still a total mystery? Even if it were true, as it is not, that we always know reason when we see it, so that some statement of what it is could be obtained from when we say it is, there would still remain untouched all the interesting questions. Most obviously the source or origin or conditions of possibility of reason; but also an explanation of the particular connection with other cognitive and behavioural phenomena and capacities which the analysis would, hopefully, make a start on mapping out. Yet how could that mapping-out be more than a start unless it were, in effect, an empirical-theoretical process of scientific enquiry?

The idea I suggest below that we should pursue consists in applying the articulated two phase analysis of evolutionary explanation to

knowledge, attempting to find different periods in the mental life of organisms which can be seen as causal and cybernetic. I argue below that this can be done, that a causal phase is seen when some cognitive adaptation, which we may think of as a rule or set of rules, is functioning properly - just as any given temperature graph could be given a causal explanation in Ch.5. But when breakdown occurs - when a rule is revised, the room gets too hot - it is necessary to appeal to the cybernetic account of how the cybernetism functions to explain why it cannot. The abandonment of rules because they won't work is selection acting against a certain kind of organism.

6.2 The duality of rational life.

My account of evolutionary explanation clearly requires that I establish a dichotomy within rational explanation which falls under the causal / cybernetic rubric. It would however be a mistake to suppose that the common contrast of theoretical and practical reason, that is of rational belief and rational action, will provide this dichotomy.

Suppose I am driving along and a swan walks into the road. I form the belief that I see a swan, and in consequence I put on the brakes. What, if anything, is rational about this mundane episode? Is my action a rational one? "Why are you braking?" - "There's a swan on the road" - "So what?" - "I don't want to hit it." If this is satisfactory, it is because braking will help prevent what I want to avoid. The means I have chosen are apt to the end I desire, and that is why I chose them.

(cf. "Why are you braking" - "There's a swan on the road" - "So what?" - "I hate swans, let's get it." Here I am rational but ill-informed about brakes perhaps. Then cf. "Why are you accelerating?" - "There's a swan on the road" - "So what?" "Swans scare me, we have to get away" - and the swan is frightened off. Here the "aggression" is apt, but only the accidental result of a botched fright. Here I am irrational, if I think to so escape). What about my belief that a swan is there? Is it rational? Certainly, if I can give a reason for it ("just like my pet Ugly"); but arguably also if I have no reason in mind and simply register it unthinkingly, as long as I can if necessary produce some kind of justification. ("Well, there are only a few kinds of large white bird, and I've only ever seen swans, and .."). This is fine even if the perception is illusory of course, if I am genuinely deceived.

What unites the two cases is that reasons can be demanded and provided; but what distinguishes them? It is fundamentally the direction of the relation between myself and my environment - in the case of the belief, the environment acts on me (largely!), for the action, vice versa. But this kind of distinction quickly becomes useless when one considers less simple cases - like someone's belief that nonsense should be exposed and his practice of continually asking sarcastic questions. However one might disentangle these two things, it doesn't look like the causal / cybernetic distinction.

I propose instead that within the classes of both beliefs and practices we may distinguish endorsements / acquisitions from repudiations; for beliefs this is obvious and for practices almost equally so. What is less clear is where simple actions like my braking fit in to this

distinction. One can repudiate a practice (like a bad habit), but surely not an action. But then one can, by disavowal : "I braked to avoid the swan, which was stupid since I knew the truck was so close behind." This seems to me to parallel the repudiation of my belief in the swan's being there, if on getting closer it turns out to be a sheep. Still, there seems to be a difference in that the action still did happen whereas the belief now no longer exists. True enough; but then the mistaken belief did exist, and the action is no longer endorsed. What would distinguish, after the event, my endorsing my past action from my endorsing my past belief? The endorsing would be the same, a kind of confident assumption or perhaps even making sure that good reasons were to be had for the belief/action respectively. Being of a belief or of an action would provide what distinction remained to the endorsements. Similarly for repudiations.

6.3 Ontology of rational life

The dichotomy alluded to in 6.2 has definite implications for the nature of a rational explanandum, and the next requirement must be some ontological underpinnings. What are we interested in explaining? Our original problem was to find a sense in which one theory is better than another which can always be applied; but this later became something more like seeking assurance that that there will always be a sense in which a successor theory is better. This is too weak, though, unless the senses in different cases are themselves systematically connected. That is what we hope the evolutionary framework to achieve. The basic notion is that the

relation in question is best explicated as is the relation of the human arm to early mammalian forelimbs, which we also regard as better - better adapted. In the latter case, this is achieved by reference to the relative fitness of organisms with the traits in question. The same must be true about theories seen as traits of organisms : that is, we are going to conceive the explanandum as the changing pattern of believings-that-X among some class of organisms, generally human ones. I repudiate such "poor man's Platonism" as Popper's third world, for being both unnecessary and mystery-mongering (how do his theories get eliminated?)

So, in the end we must eschew talk of propositions and such-like where these seem to introduce extra entities beyond the believings of organisms and the world they believe things about. I am going to assume without argument that any cognitive trait of an organism has some physiological substrate (a state or process of the nervous system, most likely). I shall argue below that they are cybernisms. Thus I am adopting a dispositional account of belief, but a stronger one than, say, Ryle's. For I assume that a belief is a capacity for believings : some genuine entity which is causally efficacious in the process of bringing a belief to conscious awareness. What belief some organism's brain-state S, say, is, is determined by its relation to, on the one hand the environment and on the other, to other organisms' similar brainstates. A proposition requires redundancy in the set of brains of its believers and disbelievers; and it also requires a redundancy between the latter and their environment. It must inform them.

These covert cognitive vehicles (CCVs) must in some structural ways be related to the overt cognitive vehicles of speech, writings, etc.,

but the details are not presently relevant. (Reasons for the existence of such relation emerge in 6.4 below). What is relevant at this point is that they must bear two kinds of relation to each other. On the one hand, relations of similarity between CCVs in different organisms constitute the identity of the abstract "entities" we are being ontologically wary about. On the other hand there must be relations between the various CCVs in a particular organism. What kind of relations?

For our purposes the most important such relations will be those which I earlier adverted to in terms of rules. A rule I characterised then as a general proposition of normative import. What this must amount to is a certain kind of relation of one to another cognitive vehicle in some specific organism - else how could the same proposition be a rule for you and not for me? Normative force in one cognitive vehicle consists in control over the functioning of others. I argue in 6.5 below that this relation of rule/ruled is an example of cybernetic control in the sense of Chapter 5.

A final preparatory remark, however, is required. Recall that there was no need in Chapter 5 to discuss the source of the energy which enabled the homeostatic room to work. Similarly, we shall quite ignore the question of just how the exercise of control is "powered" - except to say that it must clearly be done through the emotional results of social interaction. Which presupposes the publicity of rules, at least in the sense of their being conveyed though not necessarily articulated. You cannot exert normative force

purely on yourself - as of course Wittgenstein has argued somewhat differently, (the private language argument).

6.4 Causal Phases of Rational Life : Acquisitions

What is the nature of reproduction with variation in the cognitive realm? Most simply, an organism can reproduce one of its cognitive traits by telling stories about it linguistically. Much informational redundancy is created this way. This is I think undeniably a causally explicable process (though of course most of the detail is quite unknown at present). But in some manner the information in my head gets into yours via a chain of events, the middle of which is as classically causal as can be - I mean, sound waves. Moreover, there is clearly no need, in general, for any reciprocal influence of auditor on speaker (cf. radio) so that no cycle exists. (Conversation is a more complex phenomena touched on in 6.5 below).

That variation, and random variation at that, attends this process is a well-known pedagogic difficulty. But even given ideal conditions, full attention, intelligence and goodwill, still an exact reproduction is grossly unlikely for the simple reason that any cognitive attitude to some organism must be articulated on the cognitive system already present, and use mostly elements thereof. But this will be already greatly different from the sources for reasons obvious enough when we reflect on the unlimited variation in the histories of individuals and their private environments over time. And in any case the imprecision of the public language used for transmission must suppress much of the specificity of the source's covert cognitive vehicles.

There is another way in which reproduction with variation of cognitive traits such as beliefs occurs, however, and that is within the one organism. There is a kind of minimum case in which the belief-set of an organism is simply maintained from one time to another, subject to chance interferences of malfunction of the causal processes which make it possible (e.g. brain damage). This occurs for physiological traits too, but is irrelevant there because genetic reproduction is central to our explanations. Equally so for cognitive traits in that if no reproduction by communication occurs, persistence in a single organism is of ultimately no account. Opportunities for reproduction are much more ubiquitous in this latter case however, and the disentangling of reproductive and selective phases much more of a conceptual regimentation.

But there is a further case of this reproduction within the same organism which is very important, namely that in which new beliefs (or other cognitive vehicles) are simply added to the preceding system, either through perception, or invention or inference. I have an empirical leaning toward resisting the idea that the latter can be anything but a complex combination of perception, memory and chance ("lateral thinking", "bisociation", etc.), but such a doctrine need not be established here. What is worth scotching is the possible objection that physiological reproduction with variation doesn't introduce new characters, just variations on old ones. This is however only largely true, for how else could an increase in the number of traits, as has certainly occurred across species, have occurred? Certainly, much of the existing mathematics abstracts from this possibility, but

that is for mathematical reasons, not biological ones.

This point bears on one of the objections to evolutionism discussed in Chapter 4, namely the charge of non-randomness of variation. We must be careful what is meant here. Maynard Smith expresses this "central dogma" as the thesis that information flows from gene to soma but not vice versa. Now on our view this is not in any way breached by the cognitive case, for we are ignoring the genome altogether : the reproduction we are concerned with is not genetic anyway! I don't deny that variation is not random with respect to the environment, but then neither is the course of developmental physiological processes. Yet the latter are certainly part of the evolutionary explanation.

The obvious objection to this, though, is one that biologists (as opposed to philosophers!) would have made long ago : namely that precisely the difference between evolution and history is that the former depends on genetic reproduction. Of course, in a sense the latter does too, to provide human organisms, but the point is clear enough. The claim is made in different forms by Maynard Smith, ch. 19; Mayer (Boston Studies 14); Longuet Higgins (in TATB4), etc. I cannot deny that there is a clear distinction here, but I do resist the way it is used to confuse semantic and empirical issues. Let the biologists use the term evolution so, if they wish to refer only to the explanation of neoDarwinism as presently formulated. Then though they owe us a word for both what Darwin's theory was about (since, as I have mentioned, Victorian notions of reproduction are vague), and for the larger process of evolution/history which the forthcoming theory will treat. (Waddington's work on "canalization of

development' and 'genetic assimilation', theorized hopefully by Thom's theory of catastrophes, will be one extension of the theory beyond neoDarwinism. A mathematical ecology, probably systems theory based, will be another. Complete and integrated cybernetic redescription will be a third.)

As in Chapter 4, the processes discussed here will reveal cybernisms if examined at a deeper level of detail; the same kind of remarks may be made here as there.

6.5 Cybernetic phases of rational life : repudiations

6.5.1 Perception is under cybernetic control

That perception is a process under cybernetic control is a proposition widely accepted by psychologists (e.g. Gregory). Consider again my confrontation with the whitish bird on the road, and my tentative linguistic response "There's a swan on the road". This is not simply determined in a reflex way : on the contrary I would be alive to possible features of the experience which might override that response, and moreover if I am in any doubt, then the tentative response will itself result in closer attention to what I can see, which will once again rebound on the propriety or otherwise of that response. Nevertheless, if the rules governing my use of 'swan' are as elastic (plastic) as in practice they are, variations in the scene I am presented with will be suppressed by compensations in the application of the rule - "well, it could be whiter, but it is hissing!" - up to a certain point : I can recognize it as a swan despite variations in lighting, angle, condition of the bird and so on. But, there are limits of course, beyond which

the rule is breached and I withdraw the application. This is the kind of thing underlying the reluctance of some writers, as I discussed in Chapter 2 to accept the notion that language is rule-governed : since they take this to entail an iron control, that is, causal.

6.5.2 Linguistic rule is cybernetic

This case I think is typical of the way rules operate, they exert plastic control. Indeed the 3 marks of cybernetic control I suggested in Chapter 5 can all be seen to obtain in the functioning of rules. In the first place, the functioning of rules can be described so as to support strong counterfactuals. - "If the bird had been less white but known to come from that bird's egg, you would nevertheless call it a swan."

In the second place, the explanation is a capacity - the capacity to use certain words, or to recognize certain kinds of things, rather than any particular events. It is not this particular swan recognition that needs explaining but the fact that I can recognize swans.

Thirdly, the explanation, such as it is, is quite formal : just as we ignored physical facts about how thermostats and heaters work, so too here we will (indeed, our ignorance dictates it) ignore the neurophysiology which we are nevertheless convinced underlies this kind of phenomenon. What is explained is the way the vagaries of the 'environment' inform the functioning of my linguistic apparatus.

6.5.3 The last remark brings up the question of the environment of a cognitive organism. For perception in straightforward cases like the swan case, no special remarks are necessary perhaps; but it is misleading to think of the cognitive environment as being always just the same as an amoeba's would be in the same place. In a sense, when I confront an article about Quasars, my environment is physically the same as would be an amoeba's in my place, with the difference that I have an "internal environment" sharing great redundancy with what I see (so I can read it) which the amoeba has not. But a more useful description of this situation for epistemological purposes would be to regard the written report as a medium through which I "observe" features of the larger universe which I can't visually perceive directly - stars, quasars etc. There is nothing relativistic about this possibility : it is only possible to provide such a redescription in terms of "organism-believing-P in environment E 'where the description E includes facts bearing on P, because the alternative description (which we don't yet have) relating my covert cognitive vehicles to the pattern of marks on the paper is possible, which itself rests on the existence of the linguistic, cognitive community of course and moreover doesn't exist for a similarly placed amoeba.

6.5.4 With such a notion of environment, we can see what is involved in the repudiation of a belief (a set of beliefs - I will say something about theories directly). The basic schema, as Popper has been saying for half a century, is this :

- (1) A believes T
- (2) A believes $T \rightarrow 0$ (strongly)
- (3) A perceives ~ 0 (strongly)

A rejects T

A can keep to T and dismiss (2) or (3); I assume that in this case he doesn't. That is, A has some justification for accepting ($T \rightarrow 0$) and (~ 0). This requires, at least, that he has not made any inferential blunders in obtaining ($T \rightarrow 0$), nor any kind of perceptual blunder in obtaining (~ 0) and any other observations underpinning ($T \rightarrow 0$). Observations will generally be of the cognitive environment, of course, that is will require a description in line with the viewpoint of 6.5.3 rather than a Laplacean snapshot. These requirements seem both to boil down to the demand that the relevant cognitive systems be functioning properly. I have already said a little of the cybernetic nature of perceptual functioning. What about inference?

My mention of inference in discussing the addition of beliefs in 6.4 has already committed me to the view that the basic inferential step is a causal one. This requires immediate qualification, however. One obvious objective is that inferences do not occur automatically, they require effort - many people 'just can't follow' some arguments. Well, this is true, but it is also true that some inference do occur automatically, even unconsciously ("Oh, I was assuming that because ..."). Must we explain the latter kind of case in terms of some fastworking, high-effort shadow agent operating behind, but like, consciousness? If not, whatever explanation we propose

must be causal. But then why not extend it to the first kind of case, with the following rider : namely, 'unconscious' processes of inference go on within cybernisms, which exert various conflicting kinds of control over them which interfere with their 'automatic' nature. For example, cybernisms which regulate the recognition (not necessarily conscious) of an inference as being of a known valid form; cybernisms regulating the attention and hence 'energy' of the agent towards or away from various possibilities; and cybernisms of the kind illustrated by the situation we began this subsection considering; for the inferences involved in A's rejecting T, (from T and $T \rightarrow 0$ to 0; from 0 and ~ 0 to rejection of a premise) are clearly subject to control exerted by the beliefs A has in his premises : if the latter are strong enough, he won't believe the inferences valid. And yet the inferences clearly exert reciprocal control on these premises. Normally, one accepts the results of inference, but not always : and Chapter 3 argued that this might be for good reasons other than caution about mistaken inferences - one's accepted set of valid inferences may need revision. So we see that though inferences are arguably causal, they are locked into feedback cycles by the unity and coherence of the whole cognitive apparatus.

How then should we characterize A after he infers from (1) (2) and (3) that he is contradicting himself? I suggest that this is a case just like a physiological conflict - say between food-habits and one's heart-system. Something may have to give - though not necessarily, if death supervenes some other way soon enough. Similarly, one can live with contradictions by attending to other

things; but if this is not possible through the urgency of a decision for some imminent action perhaps, then something will have to give - one of the regulating circuits controlling this set of cognitive vehicles must break down. Somewhere or other a rule must be abandoned or a belief rejected. Thus is eliminated a cognitive organism of type A. Notice that this is a response to the fluctuations of the organism's cognitive environment. Our next task must be to consider such events at a population level - to consider the selection phase per se.

6.6 Statistics of rational selection

It may seem absurd to use such a phrase as 'the statistics of rational selection' : how could the rational course out of a cognitive conflict such as that of 6.5 be found by counting heads? But recall the parallel complaint about biological cases : how could the fitter organisms be determined by counting survivors without committing the absurdity of arguing in a circle? In both cases, there is an error in confusing the determination of the fitter kind of organism, in the sense what makes it so, with the determination of which are the fitter, in the sense of how we find out which they are. Our estimates of the probability distributions must be based on what happens of course, but they are only estimates. Moreover, in both kinds of case it is not impossible to arrive at estimates of an "internal" kind - for legs we can go some way on engineering considerations, for theories we have certain formal tools available. A little more on this in the next subsection.

As for standard evolutionary explanations, there are two sources

of uncertainty bearing on the functioning of some organism in its environment. On the one hand there is the imprecision attaching to the characterization of the organism (just what is a heart? Just what do I believe in the swan case?); on the other, there is the inevitable randomness of the environment - and in the case of cognitive characters such as theories this is especially evident because we know that our facts are a sample from the environment, often obtained by following some explicitly statistical procedure. Data may happen to fit a false theory, weak organisms often get lucky.

Yet, just as the statistical nature of the description of the selection phase is only apparently neutral to the underlying causes of why organism A survives and organism B doesn't, being compounded from the individual "functionalities", just so does a populational and hence statistical account of theory selection only apparently ignore the object relative merits of different theories. But since we have no independent access to the truth, we can hardly include that as a determinant except via the impact of the environment on individuals holding theories.

6.7 The Cybernetic Content Of Theories

An organism believing better theories than another can exert control as a result, in the following way. Consider A, who is in a quandary (a conflict of rules) ³ if he makes the inference from (1) - (3) to (4). Now although no-one can force A to draw the conclusion if he feels it uncongenial (which is why it is often denied that inference is a causal process), because A controls the functioning of that cognitive vehicle with others - a common ruse is to refuse

3 . cf. Hamblin (QLR)

to think about (1) - (3) by focussing attention elsewhere - nevertheless someone who has been through this and replaced T by T* which has not this conflicting consequence can present the falsity of T to A without A being able to resist (rationally) - for resistance by A will elicit the contradiction he wishes to avoid.

This is a reflection at the organism level of the relation between theories vaguely referred to as "greater content". My specific suggestion for doing something about the notoriously difficult concept of content is that a cybernetic approach be taken. A small start in this direction has already been taken by J. Greeno in Salmon ⁴, who has attempted to define the information transmitted by a theory. What is needed in our context is some measure of the relative information of one theory with respect to another.

6.8 The Evolution Of Rational Life

It should be apparent that the rationality of a given organism is hardly locatable, but if we insist on this then we might identify it with the highest level of control in its cognitive hierarchy. So for any given organism my view looks quite like the classical view : the 'reason' is a set of rules controlling its cognitive operations whose normative force arises outside the organism. But from organism to organism what it is exactly will vary; nevertheless not in an arbitrary or unconnected way.

At every level in an organism control hierarchy there is interplay

⁴. Salmon, (SRSE)
cf. also Hintikka and Suppes (II),
especially Pietarinen's article

between the ruling and the ruled cognitive vehicles, as we saw for the unfortunate A.

In judging of facts we use theories; but in judging of theories we use facts. We also use higher rules, methodological ones and logical ones. And in parallel fashion we judge them by the theories and facts they plastically control. When revisions at the highest level are precipitated by the elements of lower ones, the new rationality can be established only by its success in informing other organisms.

I shall draw some of the themes of this essay together by discussing two quite different things on which an evolutionary perspective throws some light. The first is Wittgenstein's famous and widely applauded 'family resemblance' line of attack on Augustinian semantics. This idea is "explained" in the Investigations, using the concept of a game as an example : no rule, or any other feature common to all games can be found, says Wittgenstein, yet any particular game can be connected to some others through sharing some common features. To illuminate how this is, he suggests that it is like the resemblances in a family :

"for the various resemblances between members of a family : build, features, colour of eyes, gait, temperament, etc., etc., overlap and criss-cross in the same way."⁵

I think that when Wittgenstein goes on to deride the suggestion that there is something common to all games, he goes too far. Not that there is something common to all games perhaps, but he is in danger of destroying his whole metaphor here. For a family shares a common property in the strictest material sense : any family, however extended, shares a common ancestor, and consequently genetic material which is shared by part of the family derives from a certain genotype which is also ancestral to the alternative material in the rest of the family. The many other genetic inputs change and distort the expression in different cases. This fact, that the complex pattern of overlapping similarities is the result of a causal tree, distinguishes the resemblance in a family from the similarities that might be pointed out in a set of rocks.

5. Wittgenstein, (PI), 167.

(These two contain ferrous oxide; those three copper sulphate; etc.) It is this knowledge, even in the pre-theorized form which we all have about human generation, which makes Wittgenstein's metaphor illuminating. But in consequence it commits us to expecting just such an etiological aspect in semantics. Rather than simply draw connections between gridiron and rugby in respect of shared rules, we must also pay attention to common precursors of both games in, say, Elizabethan England. Just such a historical common ancestry is what makes unobjectionable the extensions of the reference of the term 'game' which have occurred since 1700. You cannot tack on anything which shares a similarity with a game and call it a game : if some liturgical development required that in a certain service bishops should move only diagonally, we should not be tempted by the similarity with chess to dub this ceremony a game.

I am suggesting that we take Wittgenstein's metaphor seriously as semantics; and some philosophers are of course working on such a causal theory of meaning.

This essay is already long enough and an extended piece of philosophical history is not the way to finish; but some brief remarks about a particular case may make more concrete the kind of perspective whose adoption I am advocating. Consider as an example of scientific progress the advent of Darwinism itself. I think one can see that a detailed consideration would provide an example of evolutionary epistemology, and even, I suggest somewhat tentatively, of evolutionary rationality.

By the time that the joint Darwin/Wallace paper on natural selection was presented, there had been for a long time two subpopulations of cognitive organisms, namely those who did and those who did not accept the doctrine of the immutability of species (as, allegedly, laid down in the Bible). Until 1858 though, neither group seems to have had a great selective advantage over the other. But since that time, there has been a dramatic change in the populations with these traits. What is it about the theory that enables this to happen?

I think the answer to this question is that it enabled the nexus of ideas we may indicate by the slogan "like begets like" to be theoretically corrected, just as Newton's theory did Kepler's. (Like begets like is approximately true). And this in turn enabled those who revised their belief system to include the theory to exert control over those who didn't, in just the way indicated in section 6.7 - as the historical record shows by a consideration of the quality of the debates about the theory. Just this control explains why the differential in reproduction rates became so marked : the creationists convinced fewer and fewer of the uncommitted, and even lost many previous adherents, just because the theory enables

one to avoid the kind of *reductio ad absurdum* to which Philip Gosse was driven⁶ in Omphalos in suggesting that God created the world fossils and all to test our faith. The tension between this escape from threatening contradictions between the apparent age of the earth and creationist theory cannot easily resist the sneaking feeling that such a God is just too petty for belief. (Put less emotionally, the move seems enormously *ad hoc*.) It would be valuable to investigate to what extent and the ways in which the enormous public debate about Darwinism was determined by the actual theses of the theory.

One particular aspect of this particular passage of epistemological evolution is of special interest here. Although any improvement in theory must be regarded as part of the evolution of rationality - for it improves the capacities for reasoning we have, as for example in the Darwinian case just cited - most such advances do not involve any change in rationality conceived from the narrowest point of view - as the highest level of control of the cognitive system, much of which is logical. I argued in chapter three that logical revision may well sometimes be rational, but I could hardly pretend that it was very frequent. The continuance of work on evolution theory may nevertheless, indeed I think will, lead to a need for logical revision. The reason for this is the ineluctable vagueness of predicates which are needed in discussing life, as contrasted with those which physical science adopts, which are Fregean. A discussion of the logical problem I am referring to is begun in Michael Dummett's article on what he calls Wang's paradox.

6. cf. Gosse, (0) p.61

He considers the inference

- 1) 0 is a small number
- 2) if n is a small number, so is $n + 1$
- so 3) all numbers are small

The trouble here is that the premisses seem true, and the conclusion false : yet the inference is in accord with mathematical induction.

"Small" is an anthropocentric term, and here is the problem in this example. But consider this parallel case : let $S(K)$ mean "the K^{th} maternal ancestor of Coleman was human"

(where $K = 1$ refers to my mother, $K = 2$ her mother and so on. $K = 0$ gives me).

Then we have

- 1) $S(0)$ (I am human)
- 2) if $S(K)$ then $S(K+1)$ (humanity cannot have appeared from one generation to another)

yet 3) for all K , $S(K)$

must be false, if evolution is true!

This is not the place to discuss the various kinds of suggestion being made in recent literature about how to deal with vagueness, and what in particular to do about this kind of inference. Such considerations have already reached text books of logic (cf Quine, Methods of Logic, 3rd edition, p.vi - Geach's correction rests on the invalidity of an inference like the above, for

(x) ~~(y)~~ (y is the mother of x)

is false.) Whatever turns out to be the best course, which may well be, as in other such cases, to keep all options open, such development

must be counted as a refinement of our capacities for far-reaching but unobjectionable reasoning : a further adaptation of our rationality to its environment.

With these brief indications of the kind of larger project to which this essay, I think, points, I here conclude.

- (10) _____, *_____*, _____, _____.
- (11) _____, A.B., *_____*, _____, 1961, 131-4.
- (12) _____, A.B., *_____*, _____, 1961, 131-4.
- (13) _____, A.B., *_____*, _____, 1961, 131-4.
- (14) _____, A.B., *_____*, _____, 1961, 131-4.
- (15) _____, A.B., *_____*, _____, 1961, 131-4.
- (16) _____, A.B., *_____*, _____, 1961, 131-4.
- (17) _____, A.B., *_____*, _____, 1961, 131-4.
- (18) _____, A.B., *_____*, _____, 1961, 131-4.
- (19) _____, A.B., *_____*, _____, 1961, 131-4.
- (20) _____, A.B., *_____*, _____, 1961, 131-4.
- (21) _____, A.B., *_____*, _____, 1961, 131-4.
- (22) _____, A.B., *_____*, _____, 1961, 131-4.
- (23) _____, A.B., *_____*, _____, 1961, 131-4.
- (24) _____, A.B., *_____*, _____, 1961, 131-4.
- (25) _____, A.B., *_____*, _____, 1961, 131-4.
- (26) _____, A.B., *_____*, _____, 1961, 131-4.
- (27) _____, A.B., *_____*, _____, 1961, 131-4.
- (28) _____, A.B., *_____*, _____, 1961, 131-4.
- (29) _____, A.B., *_____*, _____, 1961, 131-4.
- (30) _____, A.B., *_____*, _____, 1961, 131-4.
- (31) _____, A.B., *_____*, _____, 1961, 131-4.
- (32) _____, A.B., *_____*, _____, 1961, 131-4.
- (33) _____, A.B., *_____*, _____, 1961, 131-4.
- (34) _____, A.B., *_____*, _____, 1961, 131-4.
- (35) _____, A.B., *_____*, _____, 1961, 131-4.
- (36) _____, A.B., *_____*, _____, 1961, 131-4.
- (37) _____, A.B., *_____*, _____, 1961, 131-4.
- (38) _____, A.B., *_____*, _____, 1961, 131-4.
- (39) _____, A.B., *_____*, _____, 1961, 131-4.
- (40) _____, A.B., *_____*, _____, 1961, 131-4.
- (41) _____, A.B., *_____*, _____, 1961, 131-4.
- (42) _____, A.B., *_____*, _____, 1961, 131-4.
- (43) _____, A.B., *_____*, _____, 1961, 131-4.
- (44) _____, A.B., *_____*, _____, 1961, 131-4.
- (45) _____, A.B., *_____*, _____, 1961, 131-4.
- (46) _____, A.B., *_____*, _____, 1961, 131-4.
- (47) _____, A.B., *_____*, _____, 1961, 131-4.
- (48) _____, A.B., *_____*, _____, 1961, 131-4.
- (49) _____, A.B., *_____*, _____, 1961, 131-4.
- (50) _____, A.B., *_____*, _____, 1961, 131-4.
- (51) _____, A.B., *_____*, _____, 1961, 131-4.
- (52) _____, A.B., *_____*, _____, 1961, 131-4.
- (53) _____, A.B., *_____*, _____, 1961, 131-4.
- (54) _____, A.B., *_____*, _____, 1961, 131-4.
- (55) _____, A.B., *_____*, _____, 1961, 131-4.
- (56) _____, A.B., *_____*, _____, 1961, 131-4.
- (57) _____, A.B., *_____*, _____, 1961, 131-4.
- (58) _____, A.B., *_____*, _____, 1961, 131-4.
- (59) _____, A.B., *_____*, _____, 1961, 131-4.
- (60) _____, A.B., *_____*, _____, 1961, 131-4.
- (61) _____, A.B., *_____*, _____, 1961, 131-4.
- (62) _____, A.B., *_____*, _____, 1961, 131-4.
- (63) _____, A.B., *_____*, _____, 1961, 131-4.
- (64) _____, A.B., *_____*, _____, 1961, 131-4.
- (65) _____, A.B., *_____*, _____, 1961, 131-4.
- (66) _____, A.B., *_____*, _____, 1961, 131-4.
- (67) _____, A.B., *_____*, _____, 1961, 131-4.
- (68) _____, A.B., *_____*, _____, 1961, 131-4.
- (69) _____, A.B., *_____*, _____, 1961, 131-4.
- (70) _____, A.B., *_____*, _____, 1961, 131-4.
- (71) _____, A.B., *_____*, _____, 1961, 131-4.
- (72) _____, A.B., *_____*, _____, 1961, 131-4.
- (73) _____, A.B., *_____*, _____, 1961, 131-4.
- (74) _____, A.B., *_____*, _____, 1961, 131-4.
- (75) _____, A.B., *_____*, _____, 1961, 131-4.
- (76) _____, A.B., *_____*, _____, 1961, 131-4.
- (77) _____, A.B., *_____*, _____, 1961, 131-4.
- (78) _____, A.B., *_____*, _____, 1961, 131-4.
- (79) _____, A.B., *_____*, _____, 1961, 131-4.
- (80) _____, A.B., *_____*, _____, 1961, 131-4.
- (81) _____, A.B., *_____*, _____, 1961, 131-4.
- (82) _____, A.B., *_____*, _____, 1961, 131-4.
- (83) _____, A.B., *_____*, _____, 1961, 131-4.
- (84) _____, A.B., *_____*, _____, 1961, 131-4.
- (85) _____, A.B., *_____*, _____, 1961, 131-4.
- (86) _____, A.B., *_____*, _____, 1961, 131-4.
- (87) _____, A.B., *_____*, _____, 1961, 131-4.
- (88) _____, A.B., *_____*, _____, 1961, 131-4.
- (89) _____, A.B., *_____*, _____, 1961, 131-4.
- (90) _____, A.B., *_____*, _____, 1961, 131-4.
- (91) _____, A.B., *_____*, _____, 1961, 131-4.
- (92) _____, A.B., *_____*, _____, 1961, 131-4.
- (93) _____, A.B., *_____*, _____, 1961, 131-4.
- (94) _____, A.B., *_____*, _____, 1961, 131-4.
- (95) _____, A.B., *_____*, _____, 1961, 131-4.
- (96) _____, A.B., *_____*, _____, 1961, 131-4.
- (97) _____, A.B., *_____*, _____, 1961, 131-4.
- (98) _____, A.B., *_____*, _____, 1961, 131-4.
- (99) _____, A.B., *_____*, _____, 1961, 131-4.
- (100) _____, A.B., *_____*, _____, 1961, 131-4.

BIBLIOGRAPHY

- (E) ANDERSON, A.R. and BELNAP, N.D.; Entailment, v.1 Princeton, N.J., Princeton U.P., 1975.
- (DA) ARISTOTLE; De Anima; translated by D.W. Hamlyn, Oxford, Clarendon Press, 1968.
- (DFB) ASHBY, W.R.; Design for a Brain, London, Chapman and Hall, 1960.
- (IC) _____; Introduction to Cybernetics, London, Chapman Hall, 1956.
- (TPP) BELNAP, N.D.; "Tonk, Plonk and Plink", Analysis 22, 1962, 130-4.
- (KA) BENNETT, J.; Kant's Analytic, Cambridge, C.U.P., 1966.
- (R) _____; Rationality, London, RKP, 1964.
- (MS) BUNGE, M.; The Myth of Simplicity, Englewood Cliffs, N.J. Prentice-Hall, 1963.
- (NED) BURHENN, H.; "Narrative Explanation and Redescription", Can. J. Phil. 3 (1974), 419-425.
- (BV) CAMPBELL, D.T.; "Blind variation and selective survival as a general strategy in knowledge processes" in Self-Organising Systems, ed. Yovits and Cameron, N.Y., 1960.
- (EE) _____; "Evolutionary Epistemology" in The Philosophy of Karl Popper Ed. by P.A. Schilpp, La Salle, Ill., Open Court, 1974.
- (IML) CHURCH, A.; Introduction to Mathematical Logic, Princeton, Princeton U.P., 1963.
- (RT) COHEN, J.; Review of Toulmin's Human Understanding, BJPS 24, 1973, 41-60.
- (EM) COLLINGWOOD, R.; An Essay On Metaphysics, Oxford, OUP, 1940.

- (FA) CUMMINS, R.; "Functional Analysis", JP, 72, 1975, 741-764.
- (VICS) DAVIDSON, D.; "The Very Idea of a conceptual Scheme"
Presidential address delivered to the 70th
annual Eastern meeting of the APA, Dec. 28, 1973.
- (GED) DOBZHANSKY, T.; Genetics of the Evolutionary Process, 4th ed,
New York, Columbia U.P., 1970.
- (WP) DUMMETT, M.; "Wang's Paradox", Synthese 30, 1975, 301-324.
- (AM) FEYERABEND, P.K.; Against Method, London, New Left Books, 1975.
- (MSL) FINKELSTEIN, D.; "Matter, Space and Logic" in Boston Studies V,
ed. by Cohen R. and Wartofsky M., Dordrecht,
Reidel, 1969.
- (GTNS) FISHER, R.A.; Genetical Theory of Natural Selection, Oxford,
Clarendon Press, 1930.
- (GK) GIBSON, Q.; "The Growth of Knowledge", unpublished paper.
- (MC) GINNANE W.; "Moral Concepts and Acts of Private International
Violence", unpublished paper.
- (O) GOSSE, E.; Father and Son, London, OUP, 1974.
- (AC) GOUDGE, ; The Ascent of Life, London, Allen and Unwin, 1961.
- (UN) GRENE, M.; The Understanding of Nature, Dordrecht, Reidel, 1975.
- (N) GRISS, G.; "Negationless Intuitionist Mathematics",
Indagationes Mathematicae, 6, 1944.
- (DL) HAACK, S.; Deviant Logic, London, C.U.P., 1974.
- (QLR) HAMBLIN, C .H.; "Quandaries and the logic of rules"
J. Phil. Logic, 1, 1972, 74-85.
- (RF) HARE, R.M., Freedom and Reason, Oxford, OUP, 1963.

- (LFA) HEMPEL, C.G.; "The Logic of Functional Analysis"
in Aspects of Scientific Explanation, New York,
Free Press 1965.
- (II) HINTIKKA, J.K.K. and SUPPES, P.; Information and Inference,
Dordrecht, Reidel, 1970.
- (REM) HOLLIS, M. and NELL, E.; Rational Economic Man, London, C.U.P., 1975.
- (EFL) HUGHES G.E. and LONDEY D.G. Elements of Formal Logic, London,
Methuen 1965.
- (RCT) HULL, D.; Comment on Cohen's Review of Toulmin, BJPS 25, 1974, 332-4.

- (CF) KORNER, S.; Categorical Frameworks, Oxford, Blackwell, 1970.
- (MN) KOVESI, J.; Moral Notions, London, RKP, 1967.
- (PSSR) KUHN, T.; "Postscript" in (SSR)
- (RC) _____; "Reflections on my Critics" in (CGK)
- (SSR) _____; The Structure of Scientific Revolutions, 2nd Ed., Chicago, U.C. Press, 1970.
- (PR) LAKATOS, I.; "Proofs and Refutations", BJPS, 14, 1963/4, 1-25, 120-39, 221-245, 296-320.
- (CGK) _____ and MUSGRAVE, A.E.; Criticism and the Growth of Knowledge, Cambridge, C.U.P., 1970.
- (LZ) _____ and ZAHAR, E.; "Why did Copernicus' programme supersede Ptolemy's?" in Westman, R. (ed.) The Copernican Achievement, Berkeley, U of California Press, 1976.
- (FEET) LEHMAN, H.; "On the form of explanation in evolutionary theory" Theoria, 1, 1966, 14-24.
- (GB) LEWONTIN, R.C.; Genetic Basis of Evolutionary Change, New York, Columbia U.P. 1974.
- (TML) MAKINSON, D.; Topics in Modern Logic, London, Methuen, 1973.
- (TE) MAYNARD SMITH, J.; The Theory of Evolution, 3rd ed. Harmondsworth, Penguin 1975.
- (JM) JACOB F. and MONOD, J.; "Regulatory Mechanisms in the Synthesis of Protein", J. Mol. Biol., 3, 1961, 318-56.
- (DDCS) MORTON, A.; "Denying the doctrine and changing the subject", J P, 1971, 1974, 503-510.
- (FC) MUSGRAVE, A.E.; "Falsification and its critics" in Suppes et. al. (eds.) Logic, Methodology and Philosophy of Science IV, Amsterdam, North Holland, 1973.
- (SS) NAGEL, E.; The Structure of Science, London, RKP, 1961.

- (PK) POLANYI, M.; Personal Knowledge, London, RKP, 1958.
- (AS) POPPER, K.; "The aim of Science", *Ratio*, 1, 1957.
(also in (CF)).
- (CR) _____; Conjectures and Refutations, London, RKP, 1963.
- (LSD) _____; Logic of Scientific Discovery, rev. ed., London,
Hutchinson, 1968.
- (NSD) _____; "Normal Science and its dangers" in (CGK)
- (OK) _____; Objective Knowledge, Oxford, OUP, 1972.
- (RIT) PRIOR, A.N.; "The Runabout Inference Ticket", *Analysis*, 21,
1960, pp 38 - 9.
- (ILE) PUTNAM, H.; "Is Logic Empirical"
in *Boston Studies V*, ed. Cohen R. and
Wactofsky M. Dordrecht, Reidel, 1969.
- (PL) QUINE, W.V.O.; Philosophy of Logic, Englewood Cliffs, N.J.
Prentice-Hall, 1970.
- (RR) _____; Roots of Reference, Lasalle, Ill., Open Court, 1973.
- (TDE) _____; "Two Dogmas of Empiricism" in From a Logical Point
of View, New York, Harper, 1963.
- (LR) ROSENBERG, J.F.; Linguistic Representation, Dordrecht, Reidel, 1974.
- (PB) RUSE, M.; The Philosophy of Biology, London, Hutchinson, 1973.
- (TTHC) RYLE, G.; "Thinking Thoughts and Having Concepts", in
Collected Papers, v.2, London, Hutchinson, 1971.
- (SRSE) SALMON, W.C.; Statistical Relevance and Statistical Explanation
Pittsburgh, U. of Pittsburgh press, 1971.
- (EPET) SCRIVEN, M.; "Explanation and Prediction in Evolutionary
Theory" *Science* 130 (1959) 477-82.

- (SCBS) SCHAFFNER, K. &
COHEN R.; Boston Studies 20, Dordrecht, Reidel, 1974.
- (SA) SEARLE, J.R.; Speech Acts, Cambridge, C.U.P., 1969.
- (CC) SOSA, E. (ed.); Causation and Conditionals, Oxford, OUP, 1975.
- (LSM) TARSKI, A.; Logic, Semantics, Metamathematics, Oxford,
Clarendon Press, 1956.
- (SG) WADDINGTON, C.H.; The Strategy of the Genes, London, Allen and
Unwin, 1957.
- (TATB) _____ (ed.) Towards a theoretical biology, v.4,
Edinburgh, Edinburgh U.P., 1971.
- (CS) WARNOCK, G.J.; "Concepts and Schematism", Analysis 9, 1948/9,
77-82.
- (RL) WHITELEY, C.H.; "Rules of Language" Analysis, 34, 1973, 33-39.
- (TLSF) WIMSATT, W.C.; "Teleology and the logical structure of
function statements", Stud. Hist. Phil. Sci.
3, 1972, 1-80.
- (ISS) WINCH, P.; The Idea of a Social Science, London RKP, 1958.
- (PI) WITTGENSTEIN, L.; Philosophical Investigations, Oxford, Blackwell,
1953.
- (OC) _____; On Certainty, Oxford, Blackwell, 1969.
- (F) WRIGHT, L.; "Functions", Phil. Rev. 82, 1973, 139 - 168.